

# The Near Futurist presents Dividing Lines: Can algorithms th...

📅 Thu, 7/22 4:46PM ⌚ 24:27

## SUMMARY KEYWORDS

misinformation, people, algorithms, vinay, guillaume, vaccine, debate, view, understand, facts, vaccine hesitancy, concern, trust, speak, doubt, important, genuine, question, electronic communication, information

## SPEAKERS

Vinay Nair, Guillaume Bouchard, Guy Clapperton

---



Guy Clapperton 00:04

Hello, and thanks for streaming the near future is to show presented by me Guy Clapperton, depending on starting something a little different in partnership with my new sponsor. Welcome to Dividing Lines a series of special programmes from the Near Futurist, where we will be bringing together respected innovators and thinkers to examine some of the most consequential debates in technology and society today. dividing lines is powered by diffusion and award winning international PR agency on a mission to help tech innovators to take on the status quo and transform the future faster. Oh, in case anybody's interested, they are not asking me to include any of their clients. They're leaving this fully independent and I'm very grateful for that. Vaccine hesitancy, the results of the US election the harmfulness or otherwise of 5g, these are all areas in which different areas of the internet will give you radically different pictures. You don't need to take sides to understand that someone somewhere is pumping out misinformation about a lot of subjects, and coping with that misinformation is what we're here to discuss. Today, we have two guests. First, the founder and chief executive of CheckStep. He was previously Facebook's AI integrity manager and founder of Bloomsbury AI, which was successfully exited the Facebook in 2018. checks that uses AI technology provide trust and safety solutions for social media platforms. His name is Guillaume Bouchard. Also a returning guest, CEO and founder of Lightful, who came on the show before to discuss the hesitancy

some people have about COVID-19 vaccines. His overall methods include attacking misinformation at grassroots level, which is as much about the people as anything. He is Vinay Nair. Welcome to you both. Guillaume, perhaps I can start with you. A lot of people might argue that the algorithms have created echo chambers of like minded people, they lead to polarisation and extremes. Now, one of the things you offer is, as I understand it, to throw more algorithms at the problem, is that really a good idea?



Guillaume Bouchard 02:16

Throwing algorithms is never a good idea. But we understand that algorithm just give data ability for people to deal with a lot of repetitive tasks, and to just be more efficient, because you need to be quite skilled to understand what is the truth, and to understand the context as information security. So we have multiple evidence that on some tasks, algorithms are nearly as good as human in terms of accuracy so why not use them? It makes, you know, experts or people who understand the truth much more efficient. But also there is an urgent societal crisis to solve. Yeah.



Guy Clapperton 02:59

Vinay, I don't imagine you're going to be a polar opposites in this debate. My understanding is that you're a little bit more on the people side than fighting algorithms with algorithms. Can you tell us exactly where you stand on this, please?



Vinay Nair 03:11

Yeah, certainly, it's lovely to be back guy on the podcast again. You're right, that we're not a polar opposites. I think what Guillaume says is very resonant, I think the ability of algorithms intervening to understand identify actual misinformation is very powerful. Indeed, I guess my perspective is that we need to look at it from where people are getting that misinformation, and how can we actually authentically increase trust. And ultimately, as you said, it does come down to people, I do come down on that people side. And my sense is that we kind of need to do a pincer movement here. We need to have the algorithms looking at the algorithms to understand and identify the misinformation. My view is, however, we also need to get trusted champions in communities, people who are trusted, maybe a faith leader, maybe a youth leader, maybe somebody who's active in their own community in real life, to be able to have the skills and the tools and the ability to tackle digital misinformation so that they can understand what is real, amplify that but also intervene when misinformation exists. So my sense is that it is a little bit of both, but I definitely learned a bit more on the people side.



Guy Clapperton 04:26

Guillaume, Do you have a perspective on that?



Guillaume Bouchard 04:29

Yeah, of course, it's important to say that this type of program, they don't solve really the problem of misinformation. The real problem is polarisation. And you trace it to social media. The real thing is that it enables human communities who do not have a voice today to have a voice, it's a powerful tool that the people on the side of propagating fake news are using. They're using our human biases. On these algorithms, it's a way to counteract this this huge pandemic of, of misinformation.



Guy Clapperton 05:08

That's Interesting. There are areas certainly where it's possible to have a strong view either way, we obviously had - in the last few years in the UK we had the Brexit debate where at least one view or the other was possible But there is genuine misinformation, whether it's about vaccine scares, or even who won the US election or whatever it is, there is genuine misinformation out there. What do you think it's coming from? Are we talking about malicious individuals or state actors? Do either of you have any perspective on who's putting this stuff out?



Guillaume Bouchard 05:39

So maybe I can start. So the first thing is, of course, is the natural tendency of people to believe conspiracies about something. it's naive, but here we are thinking about external foreign, often foreign organisation, such as - we often hear about Russia, but there can be a lot of other countries weaponizing information. And that's a real issue here. It needs to be to be dealt with carefully and using extremely sophisticated techniques, where sometimes genuine people are believing fake, manipulated, or just fake information. So the key point here is: Who do we trust what are the organisations that we believe are more reliable?



Guy Clapperton 06:34

That makes sense, Vinay any perspective on that?



Vinay Nair 06:37

Yeah, I think that's very sound from Guillaume. And I guess there is an aspect from the way we approach it, you know, how do we use technology as a force for good to recognise that there are those doing the exact opposite, I think the Centre for Countering Digital Hate has done some very powerful work, analysing. Now, we talked a little bit there about the state and some states, but also looking at where sometimes, you know, a dozen people can be driving hundreds of 1000s and hundreds of 1000s and millions of shares of misinformation generating 10s of millions of dollars. So that's my sense guy that we kind of need this pincer movement that we need to understand where the source of some of this is. But then a bit, as Guillaume says, there is the human aspect of why people are believing this misinformation in the first place. And my sense is that by helping to have champions who can counter that information and amplify the real, the truth, you know. You talked about vaccines, the struggle with the US election, you know, this "fake news" has sort of been used as this moniker to sometimes articulate what you disagree with, rather than what isn't true. But as you said, Brexit, where you might have had a different perspective on what we felt was right for the country. Some other aspects, it may be a lot more clear cut. So I do get the sense that, you know, that combination of state corporate and individual actors are there, we need to go after the source. But for as long as that information is out there, we need to be very, very honest, real and proactive now, in tapping,



Guy Clapperton 08:13

I'll just wondering, Vinay, your grassroots approach must be something that everyone can take part in. I mean, what should someone do if they see what they believe to be misinformation, or indeed extreme views online? Only last week in the UK, we came second in the Euro football tournament, and I refused to say anything other than came second. But the racial backlash against three of the players who well, I resist the word "missed" penalties, two of them, it was saved by the goalie he did his job. Let's not start ranting about that. But the backlash was just horrific, it may well have been a minority thing. What should people do if they see either misinformation or any hate speech like that, that are pushing us towards these polar opposites?



Vinay Nair 08:57

I think it's a great question, Guy. And I think you're right to highlight the bravery of three penalty takers, who you mentioned as well, in the final. Look, I think the grassroots approach needs a couple of different aspects to it. I've spoken a little bit already about how champions need to be able to be recognised, that they are trusted voices, and what

they need to do to be able to make sure their perspective is what dominates and what directly captures misinformation or falsehoods, in WhatsApp groups or on social media, and so on. If other members of the community or anybody frankly, on social lists seeing some of that misinformation, I guess there are different aspects: to understand what is that information? How is it making them feel sometimes? Frankly, is there even a monetary aspect so that they're being pushed down a path of clicking links, checking for veracity of quotes, checking timestamps that when that information is - you know, sometimes, I mean, Guy, these can even be satire, that you're not even quite sure if it's real or not. And so, I think being able to dispassionately look at that information should people come across it, be able to therefore try and, again speak to people who they trust to see maybe sense check the veracity of it. But I guess importantly, it's not to amplify it and share it further. I mean it in a very small, discreet and specific way. One of the I think most challenging things I find on social media, then when people do see it, you know, whether it's a quote tweet or a share of that misinformation, actually, that's something that can spread it further, which I would strongly encourage people not to,



Guy Clapperton 10:29

You're exactly right, you can amplify things if you're not careful. Guillaume, I was going to ask you how a company like CheckStep can offer help in fighting against these prejudices, these this misinformation? Can individuals get involved in that level?



Guillaume Bouchard 10:45

So there is this word, "moderation". "Moderation" also means the excess is not good. So people being able to express themselves - It's okay, but it's not okay. When the flow of bad information, so it could be racist comment, but you have to realise that this recent event about the racist remarks post the Euros, so you woke up from even yesterday or the day before about habit and behaviour. So these racist remarks are different from misinformation, There's a sense that it's not necessarily viral, is just a global, you know, rise of racism because it's easy to have strong remarks and they're highlighted by the media and by all your friends tend to be more attracted by extreme posts, but not necessarily by misinformation. So but in both cases, it's important to have this word in mind, which is moderation. Moderation means a system. We start with humans, who basically take a side and say "this goes too far". It's not... we don't want to see it on our platform. And then an algorithm such as CheckStep that provides basically the infrastructure to develop algorithms that can highlight both hate and misinformation. They can help humans to catch more, to basically moderate more on to regain a little bit of control of the content on the conversations that today is lost, because a few bad actors are getting all the attention.



Guy Clapperton 12:22

That's interesting. It also highlights the element of cultural change. I think you've touched on before, people seem more comfortable than they used to be taking extreme points of view, in that they feel that they're right or they're wrong. I just wonder whether electronic communication, a lot of it, unmoderated has consolidated that and can we actually get the oldest start of perhaps more civilised and moderate debate back?



Guillaume Bouchard 12:44

I guess we can't go back to history. So the internet is there to stay. But what we have to mention is that reporting, so users themselves, especially since last year, the George Floyd murder, there has been a lot of increase in the people reporting. You know, when you see a hateful comment, you can right click and say report to the moderator. This type of behaviour increased massively last year. So I think there is also a notion of citizenship on internet that makes people more aware and more ready to report, which enables everyone but obviously, it's not enough today.



Vinay Nair 13:25

I really welcome what you say, Guillaume. Totally agree. I think it's the ability to actually see that where the internet is here to stay. I think social media and electronic communication as you're saying, Guy, gives the opportunity for voice potentially, in a very positive way, but also allows the application of that negativity and indeed that hate. And I think the ability for individuals to actually, I think the example of George Floyd is a very powerful one. Being able to report being able to engage with the main social media, social networks and social media players, I think allows us to start mitigating some of the fact that electronic communication does make it more equitable, and give everybody a chance, but actually is giving up on something things have to align does have to still be drawn regardless.



Guy Clapperton 14:18

We've spoken a lot about deliberate misinformation so far. There's another sort of course: there's genuine mistakes and misunderstandings. I'm just wondering whether there are ways in which organisations like yours can help find that sort of error. I know a number of people are quite nervous of 5g masts, for example, they're not deliberately spreading doubt, their wish for better information is sincere.



Guillaume Bouchard 14:39

The large majority of what we call misinformation - some people call it a narrow cast misinformation, meaning a lot of people are spreading doubt because of the sorts of news that we're hearing direct from a social network. I think digital literacy can help in making people more aware of looking at the context, as we mentioned earlier, giving more, or the opposite view or maybe complimentary information for people to be able to reduce it. This basically genuine mistake, which is probably the large majority of what we see online.



Vinay Nair 15:24

I think I agree with that. My sense to your earlier question go around landing more on the people side, I do think actually, what's called the Socratic method where you ask questions to genuinely start peeling back where the source of concern is, is an important way to help understand what are often valid concerns. You know, I last spoke on your podcast on vaccine hesitancy. And my sincere view is that though there is definitely malicious misinformation that exists, a lot of hesitancy comes from fear and concern that comes from a very genuine place. And so a sense of people being listened to, people being engaged with and then getting their information when they perhaps don't trust, you know, politicians at the centre, or even in some cases, mainstream healthcare professionals. Understanding whether trust exists is the way of tackling some of that genuine misunderstanding or concern.



Guy Clapperton 16:19

I think that's right. As I think I mentioned to you last time we spoke, I have a family member who's involved in reassuring people about their vaccine hesitancy thing. And some of the community leaders she speaks to have said to her: well, you're a council officer who's going to be gone in five years time whereas they'll still be in place. And as vaccines have been around for less than a year, we objectively, we can't know what the effect will be longer term, we can only say what's likely. It may be mistaken, it may be unscientific view. But it's not hysteria. And it's not an unreasonable question to raise.



Vinay Nair 16:50

It's a perfectly valid question to raise, I think you can, you know, you did mention your family member previously. And I think the ability for them to articulate what has been done in a general sense when it has come to allowing certain medicines, vaccinations, even prior to the COVID pandemic, of what is done, how trials are conducted, the fact that they were conducted with ethnic minority groups, for example, in a safe, transparent,

efficacious way, which is not what occurred in the past, outlining what's happened with different people - whether they're of childbearing age or whether they have other concerns. Some countries, of course, have concerns for example, with the AstraZeneca Oxford vaccine, which information (through the power of the internet) flowed back into the UK, for instance. And so I think it is about being, understanding around what the concerns are. But But if you can engage and explain the importance of the counterpoints, recognise what the limitations are, and also then explain in the event, for example of not taking a vaccine, what the risks are to short term and long term health. It's not an unreasonable concern and needs to be met in an authentic way with people who genuinely are respected enough for their perspective.

G

Guillaume Bouchard 18:05

I think there is also an element which is very important for society as a whole, is allowing people to have doubt, it's very important to doubt you know, it's a sign of intelligence. But today, it seems that with this vaccine debate, we will have to be for it or against it. It's not that it's the facts that are around the vaccine, like what is the proportion of people who have had blood clots, for example, that there are facts, we often do not really rely on facts, but you know, the human biases, like we we tend to doubt this country developing this vaccine, for example. And it's very hard to capture, but a looming doubt, it's the most important thing of society. So I don't want to be, you know, anti vaccine or pro vaccine. I want both sides to express their view. But what is important is, where does it come from? Have you been manipulated in believing this or that "fact"? Just to finish on this, You know, the vaccines development took only a year. And some of the vaccines, they are based on new technologies. It's normal to have doubt. Vaccine is, again, something that requires trust, on faith in science. And it's not trivial. There are cases in the past have that led to failures, even deaths of people, and even governments hiding it, it clearly did not help. So I think there should still be a debate on we should not stop people from expressing doubts in general.

V

Vinay Nair 19:38

I agree with you. And I think one of the interesting parts of what you said is around seeing both sides of it, and how that interplays with your view of moderation. And that's where this gets complicated and intense. So, you know, I think you've framed it in a very powerful way in that people should absolutely be allowed to express that. I very much agree with And they should be met where they are, as I said, using a kind of a questioning craft method, understanding why their concerns are, and you've framed it in a 'facts versus facts' way and letting people come up with it. I think that's reasonable. But we're seeing more and more happen, for example, a lot of work happening in the philanthropy



and tech for good space around the neutrality paradox. Is it right to remain neutral? you gave the example around racism and moderation. And in some cases, you can't go outside the debate, you have to make a call. I think that's part of the challenge of where we need to see the top down aspect meet the bottom up aspect to allow debate and doubt to exist, because that stretches the intellect and it's perfectly valid. And frankly, as you say, very, very important, crucial to do the society. We shouldn't take what the government says at face value, or what companies say or even celebrities, or whatever it is at face value. But also, we need to be able to understand where sometimes a fact is very different to a view and allow space for that differential to occur as well.



Guy Clapperton 21:11

Thank you both very much. Indeed, we are now coming to the end of our time. So I think just drawing it together. much as I'd anticipated, there was a lot more in common than there were differences between you. But I think basically you're saying let's try and keep debates fact based. And let's check out facts at both the people based stuff. And of course, the algorithms are useful tools in achieving that. It's question of balance between the two. And above all, when someone disagrees with you don't necessarily treat them like an idiot. That's pretty important. So let's just round off by asking where people can find out more about yourself and your organisation. So Vinay, where can people connect with you?



Vinay Nair 21:46

Thanks, again, uy. And really, like you say, speaking with Guillaume and learning more about CheckStep has been has been great. And I think the powerful thing even though we have a slightly different perspective for a lot of common ground for us more broadly to build on so so thanks again. Yep, please, we would love people to check out our website [www.lightful.com](http://www.lightful.com) - I'm still old school and said the Ws - and follow us on social across the various platforms as well and do drop us a line and let us know.



Guy Clapperton 22:12

And Guillaume?



Guillaume Bouchard 22:13

And for us, no need for the W, just [checkstep.com](http://checkstep.com) and Yeah, it was really great to speak to about these debates. But the safety tech is a new domain in which we operate and we are

really delighted to see so much so much traction, so many platforms open to outsourcing it's a little bit like cybersecurity 10 years ago, it was used to be internal to companies are no companies open up on the accept to look at what is the best safety type solution out there. This is probably the tipping point. 2021 is a tipping point in the history of content moderation. We've survived all this regulation it that's the last one was Australia but we've heard India, and this year, UK EU on maybe in the US reform will come also this year. So in the history books, we'll probably see as 2021 months the End of the Wild Wide Web.



Guy Clapperton 23:17

We'll see about that. Maybe I'll invite you both back to discuss that again this time next year when the history books have been written! Vinay Nair of Lightful and Guillaume Bouchard of CheckStep, thank you very much for joining me. And many thanks to you for listening. That was the Near Futurist podcast with me, Guy Clapperton, and my guest in this first in the series of Dividing Lines sponsored by Diffusion PR. Don't forget to have a look at the website [nearfuturist.co.uk](https://nearfuturist.co.uk) - I'm not doing the Ws either - or my media training site at [remotemediatraining.com](https://remotemediatraining.com). I'll be back in two weeks time