

Written Representation 147

Name: Carlos Nicholas Fernandes
Technology Entrepreneur
(& retired member of the MTI's Pro-Enterprise Panel)

Received: 7 Mar 2018

Online Falsehoods - Causes, Consequences and Countermeasures

Imagine a world where the thoughts and ideas you have are not yours. They have been planted into your mind without your knowledge. The very foundations of your belief systems, your values and your opinions are not yours. They are merely artifacts of someone else's choices. We'd feel violated... if we knew. Given the choice, we would never accept it. Sadly, we weren't given the choice, and we *have* accepted it. The world I have just described is not some futuristic world from the movie "Inception". It is the world we live in today. It's the world we all live in. It's the world where the more outrageous the news, the more it is shared. The more likely it is to go viral. Truth doesn't matter. Virality does.

*"Freedom of thought... is the indispensable condition of nearly every other freedom."*¹

Fake news is not new. People have deliberately tried to spread falsehoods for centuries. However, never before in the history of mankind has the velocity of news travel been so great. The driver of that velocity is a combination of human psychology, technology and business models. It is important to understand the economics of how we got here in order to truly come up with solutions.

The Rise and Fall of Clickbait

The public have changed the way they stay informed. 20 years ago they would flip through newspapers or tune into a news channel on their televisions. 10 years ago they would visit news websites. It was only in 2011, that things began to change as social media began to go mainstream. Led by Facebook, social media had turned its attention to news. Facebook having virtually doubled its referral traffic to media sites from 12% in 2013 to 22% in 2014² had built what would become unstoppable momentum and dependency from media companies. Consumers had come to expect news to come to them instead of visiting news websites.

With significant traffic coming from Facebook, news publishers threw resources into their Facebook pages and optimized their content for social media - oftentimes even at the expense of their own websites. News publishers had inadvertently created a powerful intermediary between them and their customers. Their future was now inextricably linked to Facebook's newsfeed algorithm. Facebook and News publishers eventually embraced their respective roles: Facebook became the smart gatekeeper of content. News publishers became the dumb content pipes, constantly optimizing content for click throughs so that Facebook would deliver even more traffic to them.

¹ Palko v. State of Connecticut, 302 U.S. 319

² <https://www.inma.org/blogs/ideas/post.cfm/report-80-of-buzzfeed-traffic-driven-by-social-media>

Optimizing content for clicks placed a high burden on reporters and editorial teams, where they were judged not by the quality and rigor of their work, but the velocity with which it traveled. Highly researched and fact-driven news was expensive to create, drew less clicks and therefore was less monetizable. Facebook rewarded content that was read, liked or shared. This ushered in an era of progressively more sensationalist and exaggerated titles known as ‘clickbait’. For example, articles with titles that included, “You won’t believe what happened next...” became a Facebook staple.

The ‘experts’ had it down to a recipe - the ‘right’ headline was one that built curiosity by keeping the true topic of the article hidden until you click. People hated it, but found it irresistible to click. It worked. Upworthy, the media company that pioneered this became the fastest-growing media site of all time. Then, Facebook stopped prioritizing content based on clicks alone - instead they prioritized content based on how much time users spent on the site *after* the click. Articles with clickbait headlines (and their publishers) were punished³. In summary, stickiness was prioritized over clickiness. Publishers needed to make people linger on their articles. That had an unintended consequence.

The Rise of Fake News

Facebook killed clickbait, only to drive news publishers to search for the “next viral thing” - a new form of news - news where the articles were just as sensational as the headlines. And how do you make sensational and viral news that people that not only grabs a reader’s attention but actually triggers readership? News that is unexpected, outlandish and unbelievable. In other words, news that could simply be fake. Facebook inadvertently created an environment that fostered the substitution of something that was annoying with something that was truly dangerous and has changed the course of history.

It is therefore unsurprising that fake news ‘performed’ better than real news prior to the 2016 US presidential election - most notably the top story on Facebook was the (fake) story of Pope Francis endorsing the then Presidential candidate, Mr. Trump. The problem became widespread. According to a Pew Research Center survey, 23% say they have shared a fake news story, 14% have said they shared a story that they knew was fake at the time they shared it and 16% said they found out a story was fake after they shared it.

In a lengthy analysis, Wired Magazine outlines how a small Macedonian town of 55,000 was the registered home of 100 pro-Trump websites that were rewarded by generous Google AdSense revenue. This only happened because the fake news was shared and re-shared at a blistering pace, driving considerably more traffic than anything else would. ‘Boris’ a teen from Macedonia, apparently earned

³ <http://adage.com/article/media/facebook-publishers-clickbait-headlines/305328/>

\$16,000⁴ in ad-revenue from his two Trump websites. This worked out to be many multiples of the average monthly salary in Macedonia, which is \$371. This is yet another example where economic incentives trumped good intentions.

While Facebook has stated that it intends to stomp out fake news, it has admitted that it is not easy. Even if Facebook were to be successful, it would not be able to prevent the echo chamber effect because consumers naturally look for evidence that confirms their beliefs and read that content even more avidly. If you are optimizing for ad-revenue, there is no way to escape creating the echo chamber.

The Social Media Echo Chamber

Fake news and clickbait can be evaluated objectively. However, what if the news you see is true, but always leaning toward your existing beliefs? Indeed if the slightest non-centrist views trigger progressively more extreme articles taking advantage of the inherent human psychological weakness of confirmation bias, society gets more polarized in their views. In the end, much like the 'Blue Feed, Red Feed' experiment⁵ done by the Wall Street Journal our ideas and beliefs get skewed to result in more and more extreme positions. For example, if you were leaning ever so slightly toward Mr. Trump during the presidential campaign and you were provided with more positive views of Mr. Trump, with damning information on Ms. Clinton, you would automatically move further away from the center and closer to the most extreme supporters of Mr. Trump. So while we may not have extreme positions, social media can not just seed our minds with ideas and beliefs that we believe have come from within, but also amplify them.

Facebook is not designed to inform or educate us. Facebook is designed to engage us. According to researchers⁶ from Cambridge and Stanford University, if we have more than 10 likes on Facebook, Facebook knows you better than a co-worker and if we have more than 300 likes, Facebook knows you better than a spouse. Such a deep intimate knowledge of us can be exploited to play on our fears, exploit our weaknesses and shape our thinking - anything to drive engagement. If this seems like a deep dark dystopian view of Facebook, it is appropriate to remind us that Facebook has specifically run tests to see if they could alter the emotional state⁷ of their users by tweaking the information their users see on their newsfeed. If Facebook knows us better than a spouse, it is in a position to influence us more than a spouse can. If the information you see is directed to only one point of view, can we really have freedom of thought?

⁴ <https://www.wired.com/2017/02/veles-macedonia-fake-news/>

⁵ <http://graphics.wsj.com/blue-feed-red-feed/>

⁶ <https://www.nytimes.com/2015/01/20/science/facebook-knows-you-better-than-anyone-else.html>

⁷ [Facebook Tinkers With Users' Emotions in News Feed Experiment, Stirring Outcry](#)

Fake News: It's Motivations and Consequences

Fake news is a product of a perfect storm driven by human psychological weaknesses, personalization technology and online business models. This gives the opportunity for individuals and entities to shape public opinion at a relatively low cost, provided they can engineer virality. The fake news story of Ms. Clinton kidnapping, molesting and trafficking children in the backrooms of a pizza restaurant went viral simply because of how bizarre and outrageous it was. There was no semblance of truth to the story, and while numerous publications debunked the story, it had real world consequences - not just harassment and death threats received by the Pizzeria owner and his employees, but a man actually went there and fired three shots in the Pizzeria in an effort to 'investigate' the fake news.

It's clear that fake news can have a significant real world impact. Perhaps, more troubling however, is the fact that the economic model giving rise to fake news can be easily manipulated with goals as diverse as commercial gain, political advantage and foreign interference. Leveraging fake news to obtain an advantage in an election or seeding discord in society is a tool that can easily be used by foreign governments against Singapore. Much like how the United States' election was a victim of foreign interference, it is not a stretch of the imagination to think that the same could happen to Singapore. It's a battle that leaves the battlefield, but it's like seeding a cancer in the enemy that will eventually kill itself from within. This is, of course, extremely dangerous.

Fake news is different from propaganda, but there can be a thin line that separates the two. Propaganda has some grounding in reality, but facts are selectively chosen (or excluded) to present a skewed picture of real events. This, in some circumstances can be substantially harder to detect and while less viral, potentially even more dangerous because it could resonate with open-minded readers.

The Dark, Scary Future of Fake News - Video (or 'Deepfakes')

The recent article on Lian Wanbao where the headline was doctored from "Outdated law 'saved' the accused from harsher penalties" to read "PAP MP 'saved' the accused from harsher penalties" is emblematic of how dangerous a fake picture can be. It can shift public opinion. It can challenge the trustworthiness of our institutions - in this case our judiciary. A picture, after all, can tell a thousand words - and can consequently be a thousand times more dangerous. But what about a fake video? Fake video makes fake news even more terrifying.

Researchers at the University of Washington created a fake video of President Obama that must be seen⁸ to be believed. The fact that this is the work of researchers might give us some comfort. One

⁸ <https://www.youtube.com/watch?v=AmUC4m6w1wo>

might think that since this work is still in the research domain we have some time before it becomes mainstream and a real threat to our society. Regrettably, this belief would be misguided.

Today, readily available, free, AI technology such as FakeApp⁹ can be used to transpose a picture of a celebrity or famous person on an existing video to create a fake video - what's called 'deepfakes'. Any person can easily create the appearance of a public figure doing anything by using this technology. The New York times reports that creating a deepfake¹⁰ costed the writer less than US\$100. Granted, this commoditized fake video technology is not as advanced as the work done by the University of Washington researchers, but it comes fairly close and it will get there very quickly.

This type of fake video technology not only affects the political landscape, but has also been employed to create fake celebrity pornography. The problem became so significant, it has been banned¹¹ - even by the most permissive pornographic sites. When one thinks of the current problems with 'revenge porn' (some cases of which have occurred in Singapore) one can only imagine the trauma that fake 'revenge porn' can create.

Extortion is another consequence of fake video that cannot be understated. People who have not done things that a video purports they have may feel compelled to pay off the extortionist knowing full well that fake videos with scandalous content go viral and any subsequent clarifications and retractions would be too little too late because they get a miniscule fraction of distribution when compared to the original (fake) video/story.

How Singapore can Prevent Fake News

Many know that as technology gets more sophisticated, it gets harder to discern what is actually true. While the broad awareness of fake news is generally a good thing, it results in a deeper mistrust of anything you read or see. This mistrust can be readily capitalized upon by those with malevolent intent.

Already, there are numerous instances where 'real news' has been called 'fake' so as to diminish the impact of news that would otherwise be unfavorable. The word 'fake' is a broad brush stroke that can readily be employed today. We need to protect ourselves against instances like this. We cannot allow a total collapse of trust that builds a sense of confusion with people not knowing what to believe anymore, and being unable to be informed.

⁹ <https://www.fakeapp.org/>

¹⁰ <https://www.nytimes.com/2018/03/04/technology/fake-videos-deepfakes.html>

¹¹ <https://www.theverge.com/2018/2/7/16982046/reddit-deepfakes-ai-celebrity-face-swap-porn-community-ban>

As such, Singapore must strive to achieve three goals: First, it must eradicate fake news. Second, it must empower its citizens to get a complete picture of events and challenge the propaganda. Third, it must prevent the term “fake news” being used to discredit real, authentic news. There is no single solution to achieve all these goals.

When it comes to fake news - or news that has no semblance of reality, legislation could be a useful tool. Individuals or entities that knowingly create or distribute fake news should be punished. When it enters the realm of propaganda and where you may have online dissent or news that has been reframed to shape the point of view of a particular party (and other such grey areas), things get more tricky. The Government stands to get unfairly maligned if legislation is viewed as a tool to silence dissent. An independent ombudsman, as well as the passing of a Freedom of Information Act would help assuage this concern, but I shall not discuss this in greater detail since others have already advocated for this.

The best way to solve the problem with deliberate online falsehoods is to employ a combination of technological and community solutions, in addition to legislative ones already discussed. One of the solutions I have been exploring is the use of blockchain to transparently track the source of news, and to enable a community of fact checkers to rate the truthfulness of fake news. The Government can have representation, but so can ordinary citizens that review references and facts to assess (and rate) the truthfulness of news. Only if news reaches a certain ‘confidence level’ can it be allowed to be published.

The the above solution, however, raises another problem. How would one address breaking news? Should we withhold news until it is confirmed? What about a terrorist attack? One way to handle this is to allow news to be published if it is “breaking”, but rate it as “unverified”. Communities, police and other Government entities can then fact check and seek to verify the content and quickly approve or disapprove of the published content.

Over time, some publishers (or writers) may develop a certain reputation to be purveyors of truthful stories and those publishers may establish more trust, thereby allowing their news to reach “verified” status quicker than an someone with no track record.

Through these methods we would enable real news, that is fact checked and referenced to get out faster and build greater velocity in social networks. In fact, legislation could play a role here and distributors of news (Facebook and their ilk) could be required to distribute only news that has been fact checked or obtained a particular verification score, that’s publicly accessible through the blockchain.

The blockchain solution identified above will help solve the fake news problem, but fake video will be a much harder problem to solve. For that, not only will one need the community working to review

content as described above, but that community must be supplemented with AI tools to fight fake AI videos, with better AI detection of fake videos. It will be a technological arms race, but it will be a race we must nevertheless participate in - and win.

As such, there are technological solutions, legislative solutions and community solutions built around trust and credibility that can help solve - or at least substantially mitigate the problem of Fake News. Getting into the detailed technological descriptions are likely beyond the scope of this paper, but I would be happy to give evidence and discuss not just technological solutions but how technology, legislation and communities can work hand in hand to preserve our democracy and strengthen our society.