THE PERSONALISATION OF TARGETED CONTENT

AMBER SINHA

INTRODUCTION

At the time of writing this article, India had 500 million Internet users—over a third of its total population—making it the country with the second-largest number of Internet users after China. For the world's largest democracy, the Internet ought to be a boon. After all, Sir Tim Berners-Lee, the inventor of the World Wide Web, had envisioned the Internet as an 'open platform that allows anyone to share information, access opportunities and collaborate across geographical boundaries'. The democratisation of information it facilitated ought to have led to a more informed citizenry, but instead what we have is the complete opposite. The average digital citizen in India maintains near perpetual information illiteracy as to where he or she receives news and information from, whether or not it is true, and how it is intended to manipulate them. This is, in large part, because social media has become the primary source of information. There is a good chance that you, like me, are one of the 241 million Indians who are on Facebook. For those of us who consume and engage with information through platforms such as Facebook and Twitter, the Web has been reducing to a personalised, and therefore narrower, version of itself.

Over the last few years, the phenomenon of 'Fake News' has received significant scholarly and media attention. In March 2017, Berners-Lee had called for a crackdown on fake news, stating in an open letter that 'misinformation, or fake news, which is surprising, shocking, or designed to appeal to our biases, can spread like wildfire'.[1] Gartner, which predicts annual trends in technology,

Winter 2018–Spring 2019, Volume 45, Numbers 3 & 4

recently highlighted 'Increased Fake News' as one of its predictions for the next few years to come. The Report states that by 2022, the 'majority of individuals in mature economies will consume more false information than true information'.[2]

However, the discussion on fake news in India has remained limited to a few online stories, and a few events and workshops. Further, the discussion on fake news suffers from a lack of conceptual coherence, and this severely limits our ability to think about how it ought to be addressed. Here, we seek to understand what is meant by the term fake news, and its relevance in India. While misinformation has existed for centuries, the emergence of personalised new media presents opportunities for fake news which were hitherto unimaginable.

A BRIEF HISTORY OF MISINFORMATION

In order to understand the recent phenomenon of fake news, it is important to recognise that the problem of misinformation and propaganda has existed for a long time. Historical examples of fake news go back centuries where, prior to his coronation as the Roman Emperor, Octavian ran a disinformation campaign against Marcus Antonius to turn the Roman populace against the latter. The use of misinformation campaigns against political rivals has always existed. The advent of the printing press in the 15th century led to widespread publication. However, there were no standards of verification and journalistic ethics at the time. Pettegree writes that news reporting in the 16th and 17th centuries was full of portents of 'comets, celestial apparitions, freaks of nature and natural disasters'. As the power of misinformation at the hands of newspapers was realised, state sanctions began to emerge. In the Netherlands, in the 18th century, the legal system fined and banned publishers who put out fake reports. The World Wars saw frequent use of misinformation campaigns. The German Corpse Factory story, the propaganda machinery perpetrated by Joseph Goebbels, and the anti-Nazi propaganda by the British Information Service that was used to persuade the United States to join the war are some striking examples of fake news in this period.

In India, the immediate cause for the 1857 Mutiny was the rumour about the mixture of animal bones, particularly from cows and pigs, with the flour used in greased cartridges. Farooqui (2010)

sheds light on the strategic use of this rumour by the rebels to rouse the religious sentiments of Hindu and Muslim infantry against the British occupants. The use of pamphlets and newspapers in this campaign is instructive in the spread of printing technology for news dissemination, as well as a tool for disinformation. In India, prior to the introduction of printing technology, the concept of an authorial presence in manuscripts was limited. With a strong history of oral tradition, the manuscripts written (often accounts of religious epics and stories) were without any details about the author. The idea of a title page with bibliographical details was introduced later. Farooqui reports that a review of a representative collection of early printed books in Indian languages in the William Carey Historical Library found that most books in Indian languages, such as Sanskrit, had no title page. The practice of limited bibliographical data continued in the form of pamphlets and even newspaper reports, and one sees a correlation between the lack of author details and the use of the medium for misinformation.

Leading up to the Second World War, radio emerged as a strong medium for the dissemination of disinformation in India, and was used by the Nazis and other Axis powers. More recently, the Hindu Milk Miracle in the mid-1990s, consisting of stories of the idol of the Hindu god Ganesha drinking milk offerings, was a popular fake news phenomenon. In 2008, the rumours about the popular snack Kurkure being made out of plastic became so widespread that Pepsi, its holding company, had to publicly rebut them.

The debate on fake news often pays little heed to the semantics of the term and how we understand it. Giglietto, et al. (2016) point out the lack of conceptual uniformity in defining fake news as one of the barriers to finding solutions. Their analysis involves examining previous studies on misleading information, approaches to information theory, the concept of a hybrid media system, and the literature on diffuse newsmaking and information sharing. The authors finally evolve a model, focusing on the 'process in which misleading information spreads within the hybrid media system in the post-truth era'. Through this analysis, four types of propagations are identified, drawing a distinction between the terms misinformation and disinformation, and the intentions of the author and the propagator. According to them,

pure disinformation is when both the author and the propagator share the information with full knowledge that the information is false, whereas pure misinformation is a case of both the author and the propagator believing the information to be true when in fact it is false. Between these two spectrums lie misinformation propagated through disinformation—when the information is initially thought to be true, but is shared by a propagator knowing it to be false—and disinformation propagated through misinformation—when the author is aware of the falsity of the information, yet the propagator shares it believing it to be true.

However, this classification does not look at instances such as misleading headlines and the impersonation of sources. Based on a literary survey, this paper tries to formulate a taxonomy of fake news, which will be useful to understand and address the issue in India. At the outset, we must distinguish fake news from satire and parody, and other forms of ironic content that seek to draw attention to issues by intentionally and humorously exaggerating them, often to absurd extents. Satire and parody have long been important instruments of both entertainment and social and political commentary. Portals such as fakingnews.com and unrealtimes.com are some of the biggest aggregators of online satire, spoofs and parody in India. These kinds of materials have no intention to disinform—rather, they are important tools to highlight an existing issue by drawing attention to it by means of humour. A set of criteria is drawn up which can be used to classify news, such as its nature and kind of content, the manner in which such news is 'fake', and the medium of dissemination.

MANUFACTURED PRIMARY CONTENT

The nature of content can be broadly classified into primary and secondary content. Primary content is essentially the primary subject of a news article: for instance, statements made by persons, laws and regulations, reports and papers, raw footage or photographs featured in newspieces. Secondary content, on the other hand, is information, description or the analysis of primary content. A case of manufactured primary content would include instances where the entire premise on which an argument is based is patently false. For instance, in August 2017, Republic TV reported that electricity had been disconnected at Jama Masjid, Delhi, for non-payment of dues. This was based on the false news carried by right-wing

website Postcard News. Such instances are usually a result of poor journalism and fact-checking, where media houses do not seek sufficient verification before reporting.

DOCTORED OR MANIPULATED PRIMARY CONTENT

This is usually a case of the manipulation or editing of primary content so as to misrepresent it as something else. This form of fake news is often seen with respect to multimedia content, such as images, pictures, audios and videos, where the content is edited so as to convey a different meaning: for instance, when, during the Chennai floods in 2015, the images circulated widely on social media and even posted by the Press Information Bureau (PIB) of Prime Minister Narendra Modi conducting an aerial survey of the flood-hit city were found to be manipulated. Our limited survey of fact-checking websites suggests that these two forms of fake news tend to originate outside of traditional media, such as newspapers and television channels, and can be often sourced back to social media and WhatsApp forwards, where they may have been disseminated via a disinformation campaign. However, we see more and more such unverified stories being picked by the traditional media.

GENUINE CONTENT SHARED WITH FALSE CONTEXT

These are cases where genuine content, such as text and pictures, is shared with fallacious contexts and descriptions. Earlier this year, several dailies pointed out that an image shared by the Ministry of Home Affairs, purportedly of the floodlit India–Pakistan border, was in fact an image of the Spain–Morocco border. In this case, the image itself was not doctored or manipulated, but the accompanying information was patently false.

SELECTIVE EMPHASIS OR USE OF CONTENT

In this case, the primary content in question is itself not either false or manipulated: however, the facts, when they are reported, may be quoted out of context to the extent that, prima facie, it suggests either intent to misrepresent or negligence in reporting a comprehensive account of the facts in question. Most examples of poor coverage, especially by mainstream media, which has evolved systems of fact-checking and verification and editorial controls, would tend to fall under this category, not the previous two.

These could include the selective portrayal of information so as to completely ignore or significantly downplay pertinent facts.

In October 2017, former cricketer Rahul Dravid spoke at length at the Bangalore Literary Festival. Among other things, he spoke about the on-field aggression by cricketers and the sacking of former Indian national men's team coach, Anil Kumble. Prem Panicker, writing on this issue, stated that Dravid spoke 'with great clarity and nuance; rather than court controversy, he did his best to play them down as largely the confections of the media. There was no ambiguity whatsoever in all that he said in course of a nearly one-hour interaction'.[3] However, news reports from leading dailies, such as *Financial Express* ('Do not Follow Virat Kohli Blindly, says Rahul Dravid; Makes Big Anti-Virat Statement) and *Hindustan Times* ('Cricket Controversy: Rahul Dravid says Anil Kumble's Axing was "Unfortunate"') distorted the statements and presented them without a proper context in order to create a more sensational story. This is a commonly featured tool for clickbait journalism (web content that is aimed at generating online advertising revenue, relying on sensationalist headlines to attract click-throughs) which seeks to highlight facts or statements taken entirely out of context or framed in a disingenuous manner.

Such cases also include misleading headlines or images, where the headlines or images are not representative of the content of the story. In October 2017, Postcard News ran a story with the headline, 'Finally PM Modi has Deposited 15 lakhs in Each and Every Indian's Account'. The story cited various 'achievements' of the Union government and made vague connections, the benefit from which was somehow the equivalent of receiving ₹15 lakh from the government. This is a case of headlines not being representative of the content of a news article. The use of sensational headlines to oversell, or sometimes completely misrepresent, an article is a growing impact of clickbait journalism. This is especially apparent in the case of content providers which originated on the web, for whom the monitoring of granular user engagement and clicks is standard practice.

MISINTERPRETATION OF CONTENT

The primary difference between this category and the previous one is that it does not necessarily suggest the intentional or grossly

negligent misrepresentation of facts, but has more to do with the lack of diligence in fully understanding the issues before reportage. Such misrepresentations are often encountered while reporting in fields with specialised knowledge, and which often use jargon, such as science and technology, law, finance, among others. Earlier this year, Rethink Aadhaar, a portal by activists concerned about the Unique Identification Project in India, pointed out that newspapers have reported erroneously on a Supreme Court directive that the government could ask for Aadhaar numbers for the filing of income tax returns and applying for a PAN card. Rethink Aadhaar pointed out that what transpired were merely exchanges during a 'mentioning' of the matter before the Supreme Court, but no order to this effect was passed. Such forms of misinformation, while not suggestive of mala fide intent, can still prove to be quite dangerous in shaping erroneous opinions about important issues.

This classification, while not necessarily comprehensive, takes into account the nature of content, and the possible intent of creators and disseminators of content. These factors are instructive when thinking of ways to address fake news, as policy solutions—mandatory and voluntary—will need to consider the different forms of fake news and possible impact on them. Further, it would be unfortunate if all of the previously mentioned classifications are simply painted with the same brush under the term 'fake news', as both their 'fakeness' and the impact they have varies.

PERSONALISATION OF NEWS MEDIA

How did our online experiences become so personalised? This move by social media and search and content platforms is a response to an obvious need. On the Internet, we are constantly confronted by exponentially more information than we can digest. Former Google CEO Eric Schmidt estimates that until 2003, humans had created about five exabytes of data. It now takes us just two days to create the same amount. It is hardly any wonder that we need to rely on personalised filters to find the news we want to read, the videos we want to watch, the music we want to listen to, and the gossip we want to follow. Clearly, personalisation is sorely needed; in theory, it is not a bad thing. We don't always notice it, and it could, in the long run, appear to be inconsequential to the choices we make. But once you start to examine the myriad ways in which this personalisation

shapes our consumption, especially through targeted content, it will make you question the readiness with which we have delegated these choices to algorithms.

Imagine a personal chef who prepares meals for you everyday, based only on your cravings, and pays no heed to what your body needs to stay healthy. The menu of information served up by social media platforms is created in much the same way. Pariser's seminal book called this the shaping of our information diet. Much as our bodies crave an excess of fats, sugar and salt, our minds respond with greater appetite to 'content that is gross, violent, or sexual and gossip which is humiliating, embarrassing, or offensive' (Pariser, 2011). Algorithms play on our insatiable desires, bombarding us with sensational content, pushing us to the more polarised ends of our ideological bent. They lock us into echo-chambers, providing us with extreme versions of our beliefs. While we each experience the impact of these algorithms personally, we must also recognise how they impact society as a whole.

For a democratic society to thrive, individuals need to be active participants in discourse and not passive recipients of information. Social media platforms view us primarily as consumers, not citizens. Their single-minded drive to appeal to our basest and narrowest set of stimuli may make good business sense, but that does no favours to the cause of democracy. As citizens, we need to be exposed to more than the most agreeable or extreme form of our still evolving opinions. The signals we give to algorithms through 'likes' and clicks are often a reflection of a fleeting or tentative take on an issue. A democratic society needs media and platforms that allow us to explore different perspectives and arguments before we make up our minds. Instead, algorithms seize on our half-baked opinions and hasten their crystallisation. It is bad enough that our online selves drive this propaganda, but, lately, politically aligned actors are making creative use of such platforms to inundate us with misinformation, hate speech, and polarising content designed to manipulate.

THE CAMBRIDGE ANALYTICA DATA BREACH

To anyone privy to how social media platforms work, the use of our digital data as a political tool against us was inevitable. But it was the events surrounding British political consulting firm Cambridge

Analytica that finally brought the issue to the fore. In March 2018, several newspapers broke stories about how Cambridge Analytica had acquired the Facebook data of 87 million users, and then used it for political targeting during the 2016 presidential elections in the United States. Suggestions that this was all done with Russian involvement made the incident even more scandalous. What made matters worse, or at least puzzling, was Facebook's response—it claimed that no 'data breach' had occurred. In India, the subject of privacy had entered popular discourse recently when the Supreme Court upheld the right to privacy in 2017, and the country's many Facebook users wondered how this could have happened in the United States without a breach.

To understand that, we must go back to 2013 when Aleksandr Kogan, a psychology researcher based at the University of Cambridge, created a Facebook app featuring a personality test called 'this is your digital life' to collect user data for purported academic purposes. Facebook's Application Programming Interface (API) allowed Kogan to collect data, such as details about users' identities, their friend networks and 'likes', along with the answers to the personality test. This is our first red flag—the fact that harmless looking apps on platforms like Facebook routinely collect much more sensitive data about us than we realise.

Kogan's personality test was taken by 270,000 people. Any reasonable person would expect the collection of data to be limited to this sample. But news reports famously quoted Christopher Wylie, a past employee of Cambridge Analytica-turned-whistleblower, who made the shocking revelation that Kogan was able to collect the data of about 87 million people. That is everyone who took the test—and all their Facebook 'friends'. Kogan, or anyone else, could collect this data, all the while honouring Facebook's terms and conditions for its developers. At the time, Facebook had a feature called 'friends permission' designed to facilitate exactly this (it was rolled back in 2015). It allowed developers to access the profiles of not just the person who installs their application, but that of all their 'friends' as well. Such was Facebook's disregard for the privacy of its users that this feature was enabled by default. This meant that unless you went through your Facebook settings and opted out every single time one of your 'friends' may have taken such a test or played a game on the platform, your data was collected without your permission or

knowledge. Facebook's 'friends permission' feature is our second red flag.

The final nail in the coffin was Kogan's sale of the data he had collected to Cambridge Analytica. Bear in mind that this was the first act that went against Facebook's contractual terms and policies—the first and only action that it deemed wrong. When Facebook was informed of this unauthorised sharing, all it did was send e-mails to Kogan and Cambridge Analytica requesting them to delete this data, with little or no follow-up. Kogan and Cambridge Analytica were not suspended or banned from Facebook's platform and continued to enjoy developer privileges.

Cambridge Analytica is a subsidiary of a larger behavioural research firm called Strategic Communication Laboratories (SCL). This firm describes one of its aims as the creation of 'behaviour change through research, data, analytics, and strategy for both domestic and international government clients'. In order to achieve this, SCL builds psychographic profiles of people that it uses to target advertising and content. According to Wylie, the data bought from Kogan became the basis for voter profiles created by Cambridge Analytica for the 2016 elections in the United States. It also appeared that Cambridge Analytica's activities were linked with companies and executives associated with Russian intelligence agencies. The notion of Russian involvement in manipulating the American public to help secure the Trump presidency soon made the whole world sit up.

In the aftermath of the news reports, Mark Zuckerberg, founder and CEO of Facebook, accepted in a public statement that Facebook enabled people to log into apps and share who their friends were and some information about them. He reiterated that there was no data breach. We understand now that the collection of data by Kogan without legitimate consent was not a bug, but a feature of Facebook's platform. When Kogan collected the data of 87 million people, he was only doing what Facebook had always intended for developers like him. There had been no security lapse or breach of contractual terms. Yet, this breach has much graver implications as it involves Facebook's relationship of trust with its users. While ostensibly claiming to be concerned with user privacy, Facebook had designed its platform to ensure that users simply could not exert any meaningful control over their data. It had

effectively implemented a new form of social contract for data, where consent was assumed merely by the barest forms of participation.

THE SCIENCE BEHIND POLITICAL PROFILING

The availability of cheap and easily accessible personal data offers new opportunities for political profiling. Collecting data and analysing it to create voter profiles is suddenly very lucrative business. It is not as if the methods for creating such profiles are new; they have been around for the last century in some form or the other. But the deluge of data and its new purveyors now promise more detailed insights.

In 2017, Michal Kosinski, a researcher affiliated with Stanford University, co-authored a paper that claimed that facial recognition technology, along with deep neural networks, could be used on profile pictures uploaded on social media to predict sexual orientation. Predictably, the paper generated much controversy. It was an audacious claim which critics asserted was based on a faulty premise. Jim Halloran, Chief Digital Officer of GLAAD (Gay and Lesbian Alliance against Defamation), the world's largest LGBTQ (Lesbian, Gay, Bisexual, Transgender, Queer) media advocacy organisation, called the paper reckless and without basis. Halloran declared that technology cannot identify sexual orientation. What Kosinski's paper actually showed was that algorithms could detect a pattern in the appearance of a small subset of white gay men and lesbians on dating sites. The algorithm detected differences and similarities in facial structure, and tried to predict sexual orientation based on the assumption that gay men's faces were more feminine than straight men, and lesbian faces were more masculine than straight women. According to the paper, this finding was based on the prenatal hormone theory of sexual orientation. This theory suggests that our sexuality is, in part, determined by hormonal exposure in the womb. Kosinski's critics pointed out that factors, such as less facial hair in the case of gay male subjects, may as easily be a consequence of fashion trends and cultural norms as prenatal hormonal exposure. More importantly, critics felt the paper was dangerous and irresponsible because it could be used to support an authoritarian and brutal regime's efforts to identify and/or persecute people they believed to be homosexual. After the paper was published, Kosinski went on to claim that similar algorithms could

help measure the intelligence quotient, political orientation and criminal inclinations of people from their facial images alone. Soon Kosinski faced so much flak that he was targeted with death threats, resulting in a campus police officer being stationed outside his door.

While inferring intimate details from facial traits might seem audacious, using digital traces from social networks to do the same has gained more acceptance, and has become standard practice. Social media data is turned into actionable information for advertising and targeting by building psychometric profiles. Psychometric profiling is a process to measure and assess personality and psychology against a small number of set parametres. Most of us have taken some form of personality test, often based on the Big Five Personality Model or the Myer–Briggs questionnaire. The traditional method for conducting psychometric profiling was to carry out surveys that ask questions that can reveal aspects of the participants' psychological composition. The answers from the surveys would then be analysed to create a psychometric profile of the individual or group. However, recently, researchers like Kosinski have found that instead of conducting surveys, which are expensive and require individuals to participate actively, digital traces from social media platforms can be used to predict psychological profiles more easily and cheaply. Kosinski started out as a traditional social psychologist, trained in small-sample and questionnaire research, but was drawn to the new reality of digital data collection. The use of the digital footprint as an indicator of user attributes and preferences had been at the centre of Kosinski's research for some years. In 2013, Kosinski wrote a paper in which he analysed the Facebook 'likes' of 58,000 people, and inferred sexual orientation, race and political leanings with an accuracy range of 85–95 per cent. Six years before that, Kosinski and his frequent collaborator David Stillwell spearheaded the building of a Facebook app featuring a personality test, prosaically named 'my personality'. This app was, in fact, the precursor to Kogan's app that led to the Cambridge Analytica scandal.

In 2014, SCL courted Kosinski and Stillwell and expressed an interest in buying the dataset of the 'mypersonality' app. The researchers declined on the grounds that the data had been collected strictly for academic purposes. The firm then explored the possibility of hiring Kosinski and Stillwell to do psychometric modelling,

but the deal fell through on monetary grounds. Eventually, Kogan used his app with the understanding that he would sell the data he collected to Cambridge Analytica. This app—'this is your digital life'—is said to be inspired by Kosinski and Stillwell's app.

The theories of psychometrics which guide apps that assess personality have remained unchanged from the days of survey-based research. Most personality tests and apps are based on a psychometric model called Big Five Personality Factors. In this model, every personality is mapped across five factors—Extraversion, Neuroticism, Conscientiousness, Agreeableness and Openness. In the past century and a half, a school of psychologists has worked on the assumption that all personality traits are encoded in natural language—this theory is called the lexical hypothesis. This means that the basis for personality types is not a theoretical model but the analyses of language terms people use to describe themselves. A pioneer in this field was Sir Francis Galton. In the late 19th century, he picked up an authoritative dictionary, and began noting down words he felt were expressive of character. His exercise yielded 1,000 such words. Galton's technique was refined by others in the early 20th century, and Raymond Cattell brought up the count of trait-descriptive terms to 4,500. He later distilled these terms into 35 variables. It is these variables that were repeatedly studied and turned into the Big Five Factors.

The second dominant personality model is the Myers–Briggs Type Indicator (MBTI). Unlike the Big Five, it draws from cognitive theory. This model sees personality traits as arising from differences in how we receive and process information. Based largely on the work of psychologist Carl Jung, MBTI divides cognitive functions into eight types, depending on how we perceive and judge information. Essentially, MBTI classifies people into types, whereas the Big Five measures traits on a dimensional scale.

In both these models, the profile is intended to show how individuals may take decisions, and, consequently, how they may be influenced. Even though these profiling methods are broad-brush, machine learning promises to find correlations between 'likes' and demographic details to find patterns that represent a detailed and nuanced psychographic sketch of the individual.

The effectiveness of these methods for political micro-targeting is not a proven fact. In another paper, Kosinski, et al. (2017)

argue that tweaking advertising to psychological traits of people (again derived simply from Facebook likes) can be very effective in influencing their behaviour. Sandra Matz, one of the co-authors of the paper, has said, 'We wanted to provide some scientific evidence that psychological targeting works, to show policymakers that it works, to show people on the street that it works, and say this is what we can do simply by looking at your Facebook likes. This is the way we can influence behaviour.' This research used the data previously collected by Kosinski and Stillwell, where they inferred personality traits, such as extroversion and introversion, from Facebook likes. They used this insight to target female Facebook users with advertisements for beauty products of a particular brand. The advertisements were decided simply on the basis of whether the target was introverted or extroverted. They demonstrated that tailoring advertisements to match users' personality traits made a considerable difference to purchase, compared to users who were shown those that were mismatched. Matz claimed that this was proof that consumers respond to personalised targeting if there is a granular psychographic profile of them to guide content. This claim has limited acceptance in the academic community. There are other studies that attempt to show that personalised targeting has inadequate results, especially when used as a political tool of persuasion. While a fairly accurate profile of individuals can be built using digital traces, how effective this profile is in actually changing minds remains questionable.

At this point in time, neither researchers nor political campaigns know very much about how well targeting works at persuading voters. There is some evidence to suggest that voters rarely prefer targeted pandering to general messages. And, any form of targeted messaging runs the risk of being shown to 'mis-targeted' voters. This could actually harm the candidate, negating any positive returns from targeting. The big issue for behavioural scientists is that even though detailed profiles of voters can be created, just because the voters are understood does not mean that a persuasive message to change their views can be crafted. Even though campaigns might be unable to use targeted advertising to persuade voters to shift their loyalties, it can still be a powerful tool to contain the voters within the echo-chambers of their ideology. Already, feed algorithms of platforms such as Facebook show us content they feel we

would like. In this case, it could be political posts or advertisements and sponsored messages by political parties to whom the algorithms think we belong. Because platforms prioritise sensational content, political agents have a greater opportunity to push voters to the far end of their ideological spectrum by trapping them in a virtual world which only shows them messaging from one, and often an extreme, point of view.

THE ADVERTISEMENT NETWORK

The tight rein that platforms exert over the data they amass comes down to its monetary value. Even when web platforms provide starkly different services, issues around data, its secondary uses and how it is monetised remain the same. Google's main business model is providing answers to our questions, whereas Facebook sells itself as the platform to connect you with your friends and their interests. However, both these companies rely heavily on targeted advertising for their humongous revenues.

Alexander Nix, former CEO of Cambridge Analytica, compared data profiling to baking a cake, saying that it was the sum of its ingredients. He claimed that his company had anywhere between 3,000 to 5,000 data points on Facebook users, including voting histories, age, income, debt, hobbies, criminal histories, purchase histories, religious leanings, health concerns, gun ownership, car ownership and home ownership. Even when this data is collected from other sources, such as the household datasets built by Avneesh Rai in India, Facebook provides a unique platform to collect more granular data. More importantly, it provides the ideal stage to use these profiles to target advertising and content to users. Firms like Cambridge Analytica have found it easier to collect additional data in the United States than in Europe. In the United States, there are few regulations to prevent the secondary uses of personal data collected by private parties, and freedom of information laws allows easy access to public records maintained by the government.

In India, we face an even more relaxed state of regulation governing access to data. Unlike in the United States, there is no comprehensive data privacy law. The limited privacy-related laws that do exist are weak and barely enforced. So data collectors are free to process and share the personal data of individuals as they

deem fit. The poor interpretation of privacy exceptions in the right to information laws in India means that overzealous central and state government departments often publish spreadsheets of information, revealing personal details such as name, age, gender, address, religion, caste, Aadhaar numbers, health information and financial details. States like Andhra Pradesh and Telangana have taken to governance by dashboards and spreadsheets, feeding the gargantuan appetites of data-driven businesses. India is one of the leading emerging economies, and has a large base of online users. With more than half the country yet to access the Internet, this number is bound to grow rapidly. Add to this the fact that it has no real data privacy laws. It is therefore no surprise that it is the data industry's dream.

HOW TO RESPOND TO FAKE NEWS

The nature of responses to fake news and the use of regulatory mechanisms to address it can be broadly classified as functioning at two levels: (i) those dedicated to addressing the source and content of the news, and; (ii) those addressing the medium being deployed to disseminate such news. While the focus of the first category of efforts draws heavily from traditional forms of fact-checking and verification and their emphasis is on the greater use of these tools, the latter category looks at the peculiar ways in which the nature of the medium is an enabler in the spread of content and it can be used to control the dissemination of misinformation.

As fake news can be regulated at different levels, we have looked at Lawrence Lessig's model of four kinds of regulatory modalities. The different kinds of stakeholders involved in the creation, dissemination and receipt of misinformation are: (i) content creators; (ii) content distributors who can be categorised along the lines of traditional content distributors (newspapers, broadcasters) and digital content distributors (wikis, blogs, social media platforms, search engines, online news aggregators); (iii) norm guardians which include institutional fact-checkers, trade organisations and 'name-and-shaming' watchdogs); (iv) economic supporters (advertisers, foundations as well as consumers), and; (v) governments. Each of them is examined in some detail, along with the kind of stakeholders who are most appropriate to employ these methods.

LAW

In recent times, there have been suggestions of the state's interest in regulating fake news. Prime Minister Narendra Modi recently stated: 'The press is called the Fourth Estate. It is definitely a power, but to misuse that power is criminal.' Legal sanctions, being monopolistic and mandatory, are often undesirable, because they do not leave room to experiment with both different mechanisms and actors to solve a problem. In the context of fake news, any state regulation, too, must necessarily fall within the scheme of protection that the fundamental right to free speech and expression provides.

Defamation is one of the possible remedies used by private actors to respond to fake news when it damages their reputation. However, the inconsistent jurisprudence on free speech in India—where criminal liability for defamation is attracted at a lower threshold than civil liability, as well as the refusal to allow for an 'honest mistake' as defence against criminal defamation—renders it an ill-suited tool to address misinformation without unreasonable fetters on free speech.

Any regulations created for the purpose of combating fake news, which involve the usual state sanctions of fines, taxes and imprisonment, also face similar issues. The difficulty of defining fake news raises the risk of overbroad government regulation as well. There is the worry too that opening the door to permitting government punishment of certain kinds of public discourse would grant it too much power to control speech in areas of public concern.

The other kind of use of legal instruments to combat fake news would involve positive actions, such as creating institutions and incentives. However, these kinds of measures also raise issues of the government determining the kind of speech that is worthy, or not. Similarly, the creation of government sponsored or supported whitelists of articles or news sources run the risk of becoming a proxy for state sponsored or government approved news. The imposition of distributor liability also poses issues of curbing free speech already encountered in the context of the Internet intermediary liability.

MARKETS

Market-based solutions rely on organic changes in supply or demand, or they can be intentionally created when governments

intervene in markets to promote or discourage certain economic activity through subsidies, taxes or other incentives. Some market-based responses to misinformation seen so far include Google's decision to ban portals that publish fake news articles from using its advertising platform, Adsense. Such solutions have been found to be both over inclusive and under inclusive. Google's decision was seen to impact not only portals engaged in disinformation, but also those which engaged in political satire. Further, because such steps rely entirely on financial incentives, they completely fail to address non-monetary incentives and objectives behind fake news.

ARCHITECTURE

Lessig's model of regulatory forces views both physical and digital architecture as capable of promoting or obstructing certain values, thus providing an opportunity to structure an environment that enforces certain values—such as privacy, free expression—that we deem as socially beneficial. In the context of fake news, commentators have pointed toward the architecture of platforms such as Facebook and Twitter, which are used for the dissemination of content. The algorithms driving the Trending Topics section on Facebook, for instance, demonstrate how behaviour can be regulated through architecture. The use of selection mechanisms that promote certain stories at the expense of others play a significant role in determining what gets read and shared on a platform's digital environment.

However, such policy measures are hindered by two primary factors. First, any policy measures taken by a private company, such as Facebook, will constantly face conflict from its need to take decisions determined from the profit motive. Research has demonstrated how misinformation is often created so as to thrive in the clickbait economy, appealing to polarised opinions, which would always be to a platform's economic advantage to prioritise. Second, and perhaps even more important, because of the variety of factors at play, often the impact of such policy measures is not entirely clear to the platform itself. In 2016, a US-Senate Commerce Committee launched an inquiry into Facebook's processes to look into questions, including the supposed suppression of politically conservative content and the promotion of politically liberal content. As a result, Facebook altered the selection process for

Trending Topics to be more automated and require fewer human decisions, thus reducing the scope of intervention of human bias. However, with the reduced role of human editors, hoaxes on Facebook flourished.

Research by Conroy, et al. (2015) has examined the existing technology available in the field of 'fake news detection'. This is defined as 'the task of categorising news along a continuum of veracity, with an associated measure of certainty'. The authors identify two categories of verification—the linguistic cue approach and the network analysis approach—and suggest an alternative approach that combines the two. This will, they claim, serve as the basis for a feasible fake news detector.

ONLINE SPACES DEDICATED TO DEBUNKING FAKE NEWS

In response to the proliferation of misinformation, particularly related to politically relevant news, there have been several columns and portals dedicated to debunking fake stories, 'photos, misleading headlines and bad studies' on the Internet. One of the leading examples was Intercept, a *Washington Post* blog created by Caitlin Dewey in May 2014. In India, too, there have been numerous examples, such as altnews.in and SM HoaxSlayer. Altnews.in, run by Pratik Sinha, conducts fact-checks on stories by media organisations as well as those circulating on social media and WhatsApp. However, research has shown that these measures, while laudable and useful to the engaged audience, have limited utility. Seifert and Schwartz (2012) have argued that even after fact-checkers have busted fake news, it still has a 'continued influence effect', as it continues to influence judgements.

◆

NOTES

1. 'Three Challenges for the Web, According to its Inventor', Web Foundation. 12 March 2017. https://webfoundation.org/2017/03/web-turns-28-letter/.
2. 'Gartner Reveals Top Predictions for IT Organizations and Users in 2018 and Beyond', Gartner. 3 October 2017. https://www.gartner.com/en/newsroom/press-releases/2017-10-03-gartner-reveals-top-predictions-for-it-organizations-and-users-in-2018-and-beyond.
3. https://prempanicker.wordpress.com/2017/10/31/on-rahul-dravid-and-distortions/.

REFERENCES

Conroy, Niall J., Victoria L. Rubin and Yimin Chen. 2015. 'Automatic Deception Detection: Methods for Finding Fake News.' Conference Paper. St. Louis, US: ASIS&T.

Pariser, Eli. 2011. *The Filter Bubble: What the Internet Is Hiding From You.* New York: The Penguin Press.

Farooqui, Mahmood. 2010. *Besieged: Voices from Delhi 1857.* Delhi: Penguin, India.

Giglietto, Fabio, Laura Iannelli, Luca Rossi and Augusto Valeriani. 2016. 'Fakes, News and the Election: A New Taxonomy for the Study of Misleading Information within the Hybrid Media System.' 30 November. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2878774.

S. C. Matz, M. Kosinski, G. Nava and D. J. Stillwell. 2017. 'Psychological Targeting as an Effective Approach to Digital Mass Persuasion', *PNAS*, 28 November, 114 (48): 12714–12719.

Pettegree, Andrew. 2014. *The Invention of News: How the World Came to Know About Itself.* US: Yale University Press.

Seifert, Colleen and Norbert Schwartz. 2012. 'Misinformation and its Discounting: Continued Influence and Successful Debiasing', *Psychological Science in the Public Interest*, 13 (3): 106–31.

◆◆