

Logically.

CASE STUDY

Misinformation and the Indian Election



Table of Content

- 01 Executive Summary
- 02 Misinformation and Disinformation in India
- 03 Logically's Capabilities
- 04 Credibility Assessment
- 05 Veracity Assessment
- 06 Future in India
- 07 WhatsApp Solution for a Sharing Nation
- 08 Future in Elections
- 09 Appendix



Executive Summary

“ Problem Statement

Diverse demographics, culture and language; a troubled history of communal violence, polarised politics and sensationalist media; and a recent explosion in smartphone ownership and internet access has created a ‘fake news’ crisis in India which threatens both its democratic values and the security of its citizens.

India’s Misinformation Problem

2 million

shares of ‘Fake News’
articles during the Lok
Sabha elections

1.1 million

hours per day spent
reading ‘Fake News’ online,
excluding WhatsApp

50,000

pieces of fake news
detected during the Lok
Sabha elections

One of the unique features of India’s digital landscape is the prevalence of closed networks - ideologically homogeneous groups of individuals communicating on private platforms - in which misinformation proliferates. This poses several challenges:

- **Encryption** of private messages makes tracking and analysing the spread of information through these channels difficult.
- **Accusations of censorship** and surveillance can prevent governments from tackling misinformation propagated through private groups.
- **Ethical considerations** associated with the extraction of data from encrypted, private conversations
- **Highly Influential** means of disseminating information, with users receptive to messages which fit the common worldview of the group.
- **Speed of proliferation:** Information can reach up to 1,280 individuals at the touch of a button.



Early Solutions

Platforms



Removal of high profile accounts - Facebook removed 687 accounts linked to the INC ahead of the election and accused of spreading divisive and highly partisan content.

Funding AI research, factchecking organisations and journalistic collaborations to moderate content on the platform. While traditional fact checking has had extremely positive results on Facebook, it has been less effective on WhatsApp and other closed networks.



Fact-checking

Manual verification of high profile and suspect claims. This is a slow and unscalable industry which cannot compete with the speed at which misinformation proliferates.



Event-Based Verification Collaborations

Pop-up fact checking operation during the Mexican election in 2018 which experimented with innovative storytelling techniques to optimise its output for social media audiences to great success.

Logically: Our Approach and Capabilities

Logically has spent over two years developing solutions to detect misinformation and prevent its spread. We've developed these solutions in two areas: credibility assessment and veracity assessment.

Credibility Assessment

Deployed as part of various features within our news application, Logically's pioneering credibility assessment capabilities utilise cutting-edge content, distribution and metadata analysis to produce a real-time evaluation of content's credibility. Whilst still a best-guess approach which doesn't verify specific claims, our technology is 95% accurate, analysing every possible indicator of the credibility of a piece of text and highlighting any areas in which it falls down. As a solution to the problem of misinformation in all its forms, Logically's credibility assessment is unmatched in terms of scalability and impact.

Veracity Assessment

Human Factchecking

In addition to boasting the world's largest team of dedicated factcheckers, trained by in-house journalists to meet standards imposed by our own expectations and IFCN guidelines, Logically has developed proprietary technology and innovative methodologies to maximise the efficiency, volume and virality of our factchecks. We're committed to driving an industrial revolution in factchecking, equipping the industry with the tools it needs to provide a real solution to the misinformation crisis.

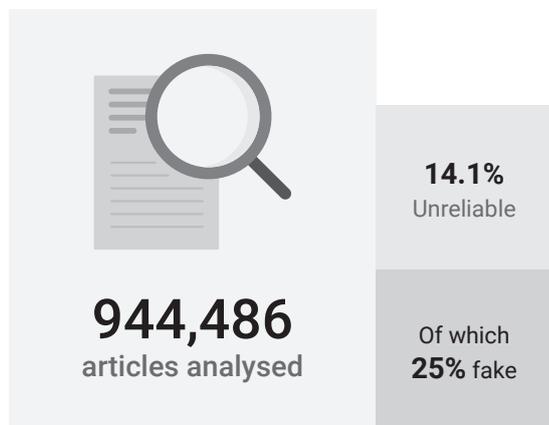
Automated Fact checking

Logically has developed machine learning and common-sense reasoning technologies that judge incoming claims against our universe of facts established by our fact checking teams, single sources of truth and trust partners. This technology is already deployed in our mobile application, where incoming claims are subject to a hybridised process which first seeks to automatically verify a claim before submitting it to our human team of factcheckers. As our knowledge of verified facts grows, so too does our ability to automatically and instantaneously verify incoming claims, making this system the first step towards the holy grail in the fight against misinformation: fully automated fact checking, providing the fast verification which is required for maximum impact.



During the 2019 Lok Sabha Elections

Credibility



Veracity



Key Findings

With encryption and privacy protection paramount, individual users will need to play a major role in any solution to the problem of misinformation, but data collected during the Lok Sabha elections highlights just how achievable a solution is. If 3 million, or just 1% of WhatsApp users, actively verify content through the platform then the country's sharing culture will help vaccinate India's WhatsApp networks against fake news. Whilst volume must increase, speed of factchecks is the vital improvement which the fact checking industry must undergo - semi-automated factchecks proved 35 times more effective than traditional factchecks in fighting fake news on WhatsApp - to debunk rumours early in their propagation path and spread fact-checks to audiences that came across the original piece of problematic content.

Moving forward amidst continued global concern around misinformation, Logically will continue to spearhead collaborative attempts to solve misinformation through our news platform and third party partnerships, underpinned by the following goals:

- Support the fact checking ecosystem with innovative new technologies.
- Support democracy during volatile election periods.
- Live fact checking for news broadcasters.
- Expose misinformation on closed networks.
- Build automated video analysis capabilities to detect doctored and 'deepfake' videos which, despite the public attention, are an almost trivial detection problem.
- Build partnerships with governments and platforms to moderate problematic content.

02

**Misinformation
and
Disinformation
in India**



**FAKE
NEWS**

SHARE

**VOTE
FOR
INDIA**

Misinformation and Disinformation in India

The prevalence of misinformation in India is worrying, though not surprising. Everything about the country's make-up has made it fertile ground for misinformation to spread. The hugely diverse demographics, the sheer number of different cultural groups and languages, its troubled history of communal violence, polarised politics and sensational media, even the geography of the country leaves it vulnerable.

A major component of Narendra Modi's 2014 election campaign was a promise to roll out internet and mobile access across huge swathes of India, and the government were hugely successful in doing so. According to the 'Household Survey on India's Citizen Environment & Consumer Economy', in 2016, while only 60% of Indian households had access to basic sanitation, 88% owned mobile phones.¹ This, compounded with the global shift to a new digital information ecosystem has contributed to the fact that India has developed what many are suggesting is a "fake news" crisis.

The effects of misinformation in India should not be underestimated. One need simply look at the widely-reported spate of mob vigilante killings that occurred over an 18 month period between mid-2018 and the start of 2019 to see the real effects of rumour and falsehood. Nicknamed the 'WhatsApp lynchings', often the killings were a direct result of rumours of child

abductions which spread over the messaging platform to rural communities. The victims were mostly strangers, passing through communities and not known to the locals who – spurred on by false rumours – carried out the attacks.

Other instances of mob violence were related to cow vigilantism directed towards Muslims and Dalits. This type of religious and communal violence often follows similar patterns to that directed at presumed child-abductors. Instigated by false rumours spread over messaging apps and directed towards someone considered 'other'.

In many instances, law enforcement were impotent and unequipped to deal with mobs stirred to violence. According to a report by the Armed Conflict Location & Event Data Project (ACLED), "the increasing number of events involving vigilante violence against presumed criminals is indicative of the perceived lack of law enforcement and lack of patience and trust in India's criminal justice system."²

The lack of trust in India's criminal justice system mirrors that of the loss of trust across most areas of the state; traditional journalism, politicians and government institutions. This atmosphere of distrust, anxiety and division set the scene for the world's largest election and was undeniably capitalised on by political actors who spread their own disinformation to leverage electoral gain.

1. Moonyati Mohd Yatid, 'Truth Tampering Through Social Media: Malaysia's Approach in Fighting Disinformation and Misinformation, The Indonesian Journal of Southeast Asian Studies, vol. 2, no. 2, Jan 2019, (pp. 203-230), p. 206.

2. The Complexity of Disorder in India, (p. 3) ACLEDdata, <https://www.acleddata.com/2018/08/17/the-complexity-of-disorder-in-india/>

The organic spread of rumours occurring on messaging apps such as WhatsApp and Telegram mimics the way political disinformation is spread by bad actors on the same platforms and can have similar outcomes. Senior Fellow at the Observer

Research Foundation, Maya Mirchandani suggests that “in India, these spaces provide both tacit and overt sanction for rising incidents of majoritarian violence as identity-based, populist politics dominate the country’s landscape.”³



3. Digital Hatred, Real Violence: Majoritarian Radicalisation and Social Media in India, ORF Occasional Paper, August 2018 (p.1).

General Elections in India

Indian general elections – overseen by the Electoral Commission of India – occur every five years. All 552 seats of the Lok Sabha are up for election. 530 members of parliament are elected from single-member constituencies in 29 states. 20 are elected to represent the 6 Union Territories and a further 2 members are elected by the President of India to represent the Anglo-Indian Community. The leader of the party which commands the majority of the house is appointed Prime Minister of India.

The sheer size of India both geographically and in terms of its population makes the logistics of a general election a mammoth task. As such, elections span a number of weeks with different regions taking turns to cast their votes before the result is finally announced. The 2019 Indian general election was held in seven phases over five weeks, taking place between 11th April and 19th May. On 23rd May Narendra Modi's Bharatiya Janata Party (BJP) was announced victorious having secured 303 seats in the Lok Sabha - India's lower house.

Historically turnout in Indian general elections has been consistently around 55% however, this has been rising over the last two elections with the highest ever turnout recorded during the 2019 election at 67.4%, accounting for over 600 million voters. This election also saw the highest participation of female voters.



Indian General Election 2019: The Campaign

The 2019 Indian General Election was announced on 10th March, though campaigning had arguably been going on for months preceding this, both overtly and covertly using social media and WhatsApp.

Prominent campaigning issues included the usual and expected debates around economic policy and national security. The BJP's demonetisation experiment and new Goods and Services Tax (GST) were both highly contentious issues. While the BJP paraded what they claimed had been a wildly successful policy at reducing corruption and the spread of black money, opposition parties argued had seriously affected small business, farmers and labourers.

Following the Pulwama terror attack in February 2019, the BJP-led NDA government conducted tactical airstrikes on locations in Pakistani territory. The events surrounding both the attack and subsequent inflaming of tensions between India and Pakistan saw a spike in misinformation, leading Trushar Barot, the head of Facebook's anti-disinformation efforts in India to tweet "I've never seen anything like this before – the scale of fake content circulating on one story." While some of this misinformation spread organically, it is likely that some were also deliberate disinformation spread by political actors to stoke division and push a nationalist agenda that would come to be helpful in the polls.

Investigations by the New York Times and the Wall Street Journal both uncovered evidence that parties on all sides made a concerted effort to spread misinformation about the Pulwama attack as well as on other issues that would benefit them electorally.⁴⁵

The major parties were hardly secretive about how they intend to use WhatsApp to spread campaigning messages. The ruling BJP was widely reported to be amassing an army of volunteers to carry out their WhatsApp campaign for the election a month or so before it was even called. They were rumoured to have taken on around 900,000 campaign volunteers specifically to work on what WhatsApp called "cell phone Pramukhs". That is nearly one volunteer for each polling station in India.⁶ The Indian National Congress (INC) – the country's main opposition party – attempted to replicate this campaigning infrastructure, though reports suggested they were some way off.

While there is nothing intrinsically illegal or wrong about using technology to spread political messages, the potential for this to be abused was evident. Indeed, The Electoral Commission of India warned parties that their social media accounts must comply with "acceptable behaviour", in line with the Model Code of Conduct that applies to traditional campaigning methods during elections in the country. Platforms themselves also

4. <https://www.nytimes.com/2019/04/01/technology/india-elections-facebook.html>

5. <https://www.wsj.com/articles/facebook-removes-hundreds-of-fake-accounts-ahead-of-indian-elections-11554129628>

6. <https://www.hindustantimes.com/india-news/bjp-plans-a-whatsapp-campaign-for-2019-lok-sabha-election/story-IHQBYbxwXHaChc7Akk6hcl.html>



acknowledged the potential for manipulation by political actors. Carl Woog, Head of Communications at WhatsApp, warned parties not to “misuse” the messaging service and insisted that it is not a “broadcast platform.” He said that any evidence of parties abusing the service could result in bans.⁷

Whilst in the run-up to the election, and during the campaign itself a number of politically-aligned social media and WhatsApp groups were removed, there is ample evidence that during the election all sides were using WhatsApp to spread highly divisive, manipulated, or completely false information.

A study by the Oxford Internet Institute (OII) discovered that more than a quarter of content shared by the BJP was ‘junk news’ as well as a fifth of content shared by the INC. In addition to this, a sample of images being shared on WhatsApp by the parties were deemed ‘divisive and conspiratorial’ and that overall “the proportion of polarizing political news and information in circulation over social media in India is worse than all of the other country case studies we have analysed, except the US Presidential election in 2016.”⁸

7. <https://www.bbc.co.uk/news/world-asia-india-47521269>

8. ‘News and Information over Facebook and WhatsApp during the Indian Election Campaign’, COMPROP DATA MEMO 2019.2, Oxford Internet Institute, May 13, 2019.

Closed Networks for Disinformation

One significant challenge for those aiming to combat the spread of misinformation is that many of the more dangerous stories and rumours are spread on social media and more specifically, on closed networks and messaging apps. The fact that these platforms are the primary tool by which misinformation seems to be spread raises a number of complicated issues.

First and foremost there are practical hurdles blocking attempts to research and counter the spread of misinformation on platforms such as WhatsApp. Because these messaging apps are mostly encrypted it is very difficult to assess the proliferation of the information that is circulating inside private groups. Unless we can gain access to groups on these platforms it is near impossible to see what content is being shared and stop anything considered false or misleading. Issues of censorship and surveillance are also raised at the thought of the government or others gaining access to private chat groups – even if the honest intention is to combat the spread of misinformation.

Another problem for researchers are the ethical considerations of extracting data from encrypted conversations between private individuals. This has already had implications on previous attempts to study misinformation in India and other countries. The OII struggled with these very issues during a study conducted on misinformation

on WhatsApp in India; “the encrypted nature of the platform (WhatsApp), amorphous structure of public groups and our strict ethical considerations pose significant challenges to joining and extracting data at scale from WhatsApp groups. We note that, our strategy does not ensure that all WhatsApp groups are adequately represented, however, forwarded content circulating within these groups could readily be shared in other private groups that the members belong to and therefore it is possible that the content analysed in this study has been viewed by a much larger network of WhatsApp users.”⁹

The closed nature of messaging platforms has also been observed to create an environment in which users feel freer to share questionable content as well as for such content to be readily believed. In his study of fake news in Asia, Andy Yee suggests that these apps create “walled gardens” in which the information shared “can especially resonate since receivers are more likely to trust their circles of like-minded people.”¹⁰

These closed networks pose significant problems for anyone wishing to tackle misinformation. Stringent research ethics stunt attempts to analyse the extent of the problem, questions of censorship and surveillance plague attempts at intervention, and the very private and personal nature of messaging groups provides the perfect conditions for misinformation to spread unchecked.

9. ‘News and Information over Facebook and WhatsApp during the Indian Election Campaign’, COMPROP DATA MEMO 2019.2, Oxford Internet Institute, May 13, 2019, p. 7.

10. Andy Yee, ‘Post-Truth Politics and Fake News in Asia’, *Global Asia*, vol. 12, No. 2, 2017, (pp. 66-71), p. 68.

Scale, Prevalence and Complexity of the Problem

To suggest that misinformation is a new phenomenon is clearly incorrect. The recent spike in interest in ‘fake news’, inspired by its popularisation and politicisation firstly in the US and then globally has, however, sparked a renewal of interest in the phenomena. In addition to this, evidence of a sustained Russian disinformation campaign during the 2016 presidential election – mostly played out over social media platforms – has woken up the international community to the fact that technology has created new avenues for bad actors to exploit and weaponise information.

In a country in which the number of voters with access to a smartphone—and by extension digital messaging apps—has nearly doubled from 21% in 2014 to 39% in 2019, it’s easy to see why such focus was paid to digital campaigning techniques by the main political parties. With the ability for a single person to share a message or story with around 1,280 different individuals in seconds at almost no cost, WhatsApp and other messaging services became key to the parties’ campaign strategies. It was reported that the BJP had enlisted nearly 900,000 volunteers to carry out their WhatsApp campaign. Other parties’ followed suit and although it is impossible to say just how effective digital strategies were in comparison to traditional election campaigning, the scale and prevalence of political messaging over digital platforms cannot be understated.

Political parties were warned in advance of the election by the Electoral Commission and by social media platforms themselves, not to misuse social media during the election. At one stage Facebook removed 687 accounts linked to the INC which they believed were being used to spread highly partisan and divisive content.¹¹ Although Congress claimed that none of their official party accounts were affected it is clear that Facebook was being used by supporters, if not the party itself, to spread misinformation and politically divisive messages for electoral gain.

While technological advancements have increased the scale and reach of misinformation, and also added to difficulties in tracking and stopping its spread, the phenomenon itself is complex and has deep roots in society. In India, a country with a history of political and religious division, much of the political messaging that is found to be problematic, touches on pre-existing cultural divisions. Benson Rajan of Christ University, Bangalore, has shown that the particular brand of Hindu nationalistic misinformation spread by the BJP in the 2019 general election can be found in its nascent stages as far back as 2014. During the 2014 Indian general election – which saw the BJP sweep to a historic victory – there is evidence of divisive and distinctly religious misinformation being spread online.¹² The apparent electoral success of these messages was consequently built on and

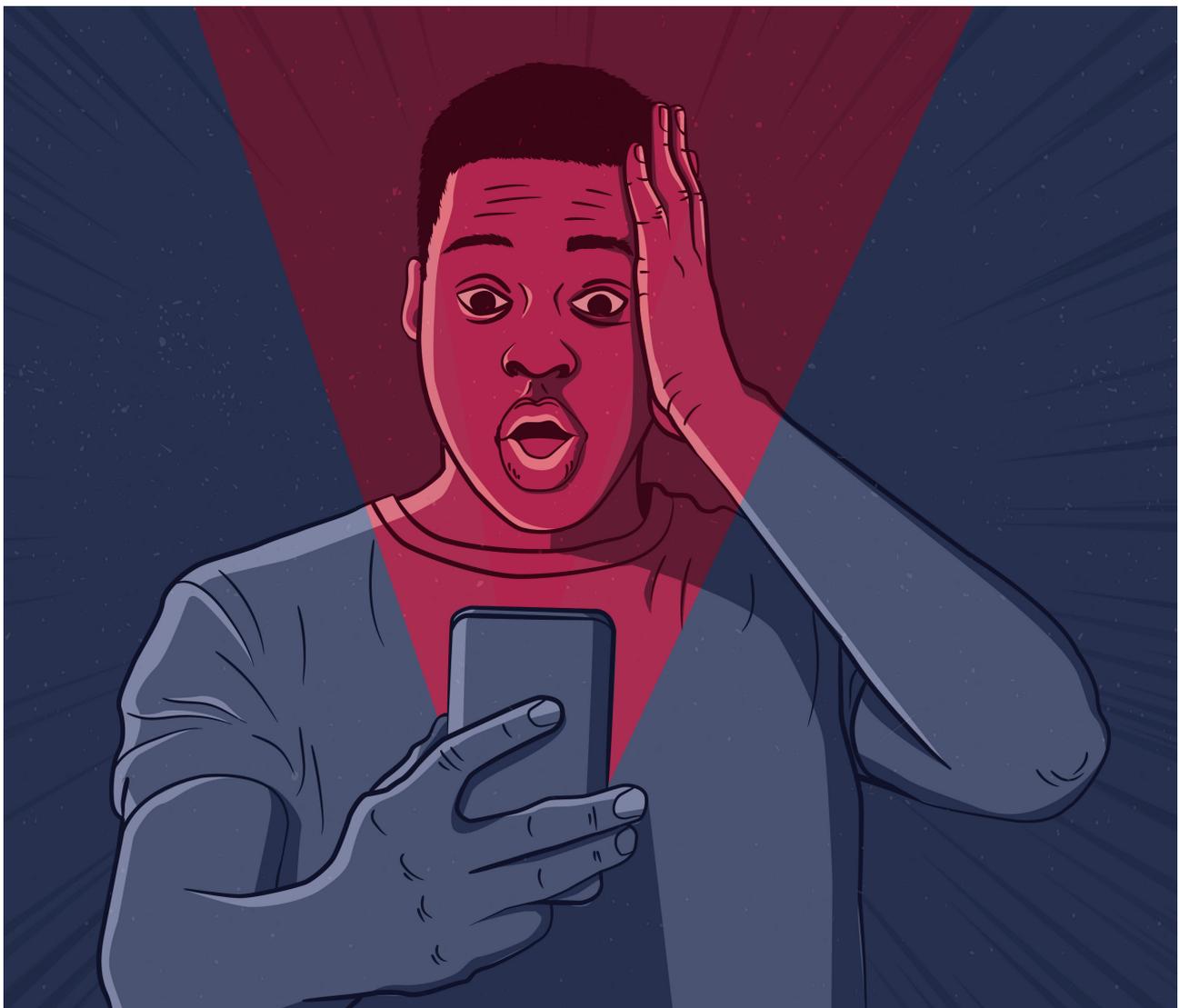
11. <https://www.wsj.com/articles/facebook-removes-hundreds-of-fake-accounts-ahead-of-indian-elections-11554129628>

12. Benson Rajan, “New Mythologies of Fake News: WhatsApp and Misrepresented Scientific Achievements of Ancient India”, Handbook of Research on Deception, Fake News and Misinformation Online, Christ University, India, 2019

perfected in time for 2019.

The blurring of political, religious and cultural messaging adds further layers of complexity to India's misinformation problem. It means that the solution is not just technological and requires us to look past the knee-jerk application of blame that has been evident in recent years. Technology can be exploited to spread divisive messages and false content. However, the readiness for society at large to accept these messages requires a far more nuanced approach. This case is made in a study released by the Berkman Klein Centre for Internet and Society at Harvard University in 2018. The authors write, "Specific technologies, under specific institutional and

cultural conditions, can certainly contribute to epistemic crisis. . . WhatsApp in India, suggest(s) that technology, in interaction with particular political-institutional conditions, can become the critical ingredient that tips societies into instability... But we have not seen sufficient evidence to support the proposition that social media, or the internet, or technology itself can be a sufficient cause for democratic destabilization at the national scale...it is only where the underlying institutional and political-cultural fabric is frayed that technology can exacerbate existing problems and dynamics to the point of crisis."¹³



13. Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics, Oxford Scholarship Online, October 2018

Platform impact

The fact that solving the misinformation crisis will require more than technological solutions and regulation does not change the fact that it has had, and will continue to have a significant impact on social media companies. Companies such as Facebook, which also owns WhatsApp, and Google, the parent company of Youtube, have already recognised the damage to their brands' reputations by their association with scandals surrounding 'Fake News'. It is not only reputational damage that these companies are risking but also legal and financial. Indeed, this year Facebook was fined \$5bn following investigations into the Cambridge Analytica Scandal and their misuse of users' personal data.

In response, the platforms have made efforts to mitigate the spread of misinformation using their products. In some cases, they have removed high profile figures from their platforms who are widely considered to spread misinformation or divisive content. Alex Jones of Infowars fame, for example, has had profiles suspended and content removed from Facebook, Apple, Youtube and Spotify. Simply removing content and users who are considered to be misusing platforms has not however proved to be a sustainable or effective method for platforms to deal with the issues of misinformation. In addition to the cries of censorship which are levelled whenever the leading social media platforms resort to bans and suspensions, the simple fact is that removing a few high-profile figures or closing down groups does not address the root cause of misinformation being spread

using their platforms.

In India, similar attempts were made to remove groups accused of spreading misinformation and harmful content in the run-up to the election. Facebook claimed to have removed hundreds of such groups – some of which they traced to Indian political parties and even the Pakistani military's public relations wing.¹⁴ Whilst these efforts should not be disparaged, they barely touch the surface of what it will take to completely solve these issues.

A number of these companies have also begun to spend large amounts of money sponsoring and financing both technological and journalistic efforts to combat misinformation on their platforms. Fact-checking organisations and journalistic collaborations created with the explicit purpose of monitoring and debunking misinformation in India and elsewhere often receive large amounts of funding from the social media giants. It is clear that the problems caused by misinformation for these companies is not just bad for their reputation but also taxing on their financial resources.

A final point to be raised when considering the impact that the misinformation crisis is having on social media platforms is the risk of falling on the wrong side of regulations. As more is understood about the effects that misinformation is having on democratic processes around the world, governments are seeking to take action themselves. Whilst the number of governments making efforts to solve these problems is on the rise, their actions vary dramatically from court rulings to

14. <https://www.wsj.com/articles/facebook-removes-hundreds-of-fake-accounts-ahead-of-indian-elections-11554129628>

the introduction of new laws and even internet shutdowns.¹⁵ Reports in 2018 suggested that India had resorted to literally turning off the internet hundreds of times over the course of a year to quell rumours circulating on WhatsApp.¹⁶ In April 2018, legislation was proposed that would temporarily or permanently suspend the accreditation of journalists suspected of spreading fake news. It was withdrawn within 24 hrs, after an outcry about its implications for freedom of speech.

The more problems we see arising from misinformation, especially problems that are seen to affect the very functions of government and civil society, the more likely it is that legislation and regulations will be introduced. The chances of social media platforms falling on the wrong side of such regulations are high and has already occurred in some countries.¹⁷



15. <https://www.poynter.org/ifcn/anti-misinformation-actions/#india>

16. <https://www.wired.co.uk/article/whatsapp-web-internet-shutdown-india-turn-off>

17. <https://www.poynter.org/ifcn/anti-misinformation-actions/#india>

Early Solutions and Fact-Checking in India

Having looked at the ways in which both the Indian government and institutions such as the electoral commission and the social media giants have tried to address the misinformation crisis without much success it would be prudent to touch on other on-going attempts. Over the last few years India has seen an explosion in the number of fact-checking operations popping up. Some are small and independent, others born out of the operations of existing publishing giants.

Whilst the fact-checking industry is growing and becoming more prominent there are a number of issues that these organisations are encountering. First and foremost there is a serious lack of funding for the organisations that are appearing in what in many ways is still a cottage industry. Despite the increasing attention being shown to fact-checkers, many still rely on volunteers, handouts, and access to open-source digital tools, which can often be rudimentary.

Some funding has been made available from social media companies, while Facebook runs a third-party fact-checking operation which sees select organisations work with the social media giant to verify content on its platforms. This scheme, however, is far from perfect. Firstly there is an issue of scale. No matter how many organisations Facebook partners with

and even with their own content moderation teams, they will never be able to fact-check and debunk every piece of misinformation on their platform. Further to this, the posts that are fact-checked are not removed from the site, only have their visibility reduced to users. Questions have also been raised over a number of organisations who have partnered with Facebook in India, with claims that they themselves have spread misinformation.¹⁸

During the election itself, an independent group of journalists, fact-checkers, and technologists worked with WhatsApp to fact-check rumours and posts on the messaging platform. Billed as a 'research project', Project Checkpoint set up a WhatsApp hotline to which users could send requests for the team to verify.¹⁹ This project too suffered from issues of scale and as it was only a research project, intended to build a database of India-specific misinformation, it likely had little tangible effect or outcome.^{20,21}

Although there have been a number of attempts at finding solutions to the spread of misinformation in India, none yet have appeared to be either successful or sustainable.

18. <https://www.nytimes.com/2019/04/01/technology/india-elections-facebook.html>

19. <https://medium.com/@meedan/press-release-new-whatsapp-tip-line-launched-to-understand-and-respond-to-misinformation-during-f4fce616adf4>

20. <https://www.checkpoint.pro.to/>

21. <https://www.poynter.org/fact-checking/2019/is-whatsapps-new-tip-line-debunking-hoaxes-about-the-indian-election-not-really/>

03

Logically's Capabilities

Over the last few years Logically has invested significant resources into developing human-machine solutions to detect misinformation and decelerate its proliferation.



Human Factchecking Team

Logically's dedicated fact-checking team, one of the largest of its kind in the world, is made up of 27 people based across India and the United Kingdom, who bring diverse skills and experience from fields that include journalism, commerce, science, politics, literature and computer science.

Each member of the team has undergone extensive, industry-standard training in fact-checking and verification practices, and specialise in the investigation of political and business claims. The methodology that they apply has been developed in line with International Fact-Checking Network specifications.

The team's work undergoes extensive internal review before publication, to ensure that it's of the highest standard. This involves multiple levels of oversight by our team of moderators, as well as final sign-off by our

senior supervisors. Everything is double and triple-checked before we send it out. We do this using a proprietary dashboard that's been developed in conjunction with Logically's internal product development team, and is optimised especially for our fact-checking work.

The team continues to expand, with an additional twenty fact-checkers coming on board in Fall 2019 to enable 24x7 operation. Since the beginning of Logically's fact-checking work for the 2019 Lok Sabha elections in India, we've published over 6000 claims, and aim to complete between 100 to 300 fact-checks in a day. Looking forward, we're preparing to play an active role in addressing information disorders in the UK and US, ahead of their respective elections.



Automation to Augment Value

At the heart of our approach is the principle of extended intelligence [see later]. In the current environment of information warfare, bad actors are equipped with sophisticated tools and bots. By ignoring potential assistive applications of AI technologies we risk bringing a knife to a gunfight on the information battlefield.

Our view of our role in the fact-checking ecosystem is to bring the industrial revolution to an industry described by many as a cottage industry.

As well as scaling our human fact-checking team and equipping them with cutting-edge dashboards and workflows to improve efficiency in line with a constantly expanding user-base, we're also developing pioneering automation technologies. Together, these technologies will stimulate a technological revolution in the fact-checking industry. Just as initially the assembly line and more recently the robotic arm enabled the rapid development of the automobile into the everyday product it is today, Logically's fact-checking dashboards and automated veracity assessment tools will enable the publication of fact-checks on a scale previously unimaginable.

In addition to the scalability of automation, the lower latency of automated fact-checks and verifications makes them orders of magnitude more effective in countering mis/disinformation compared to their traditional counterparts.

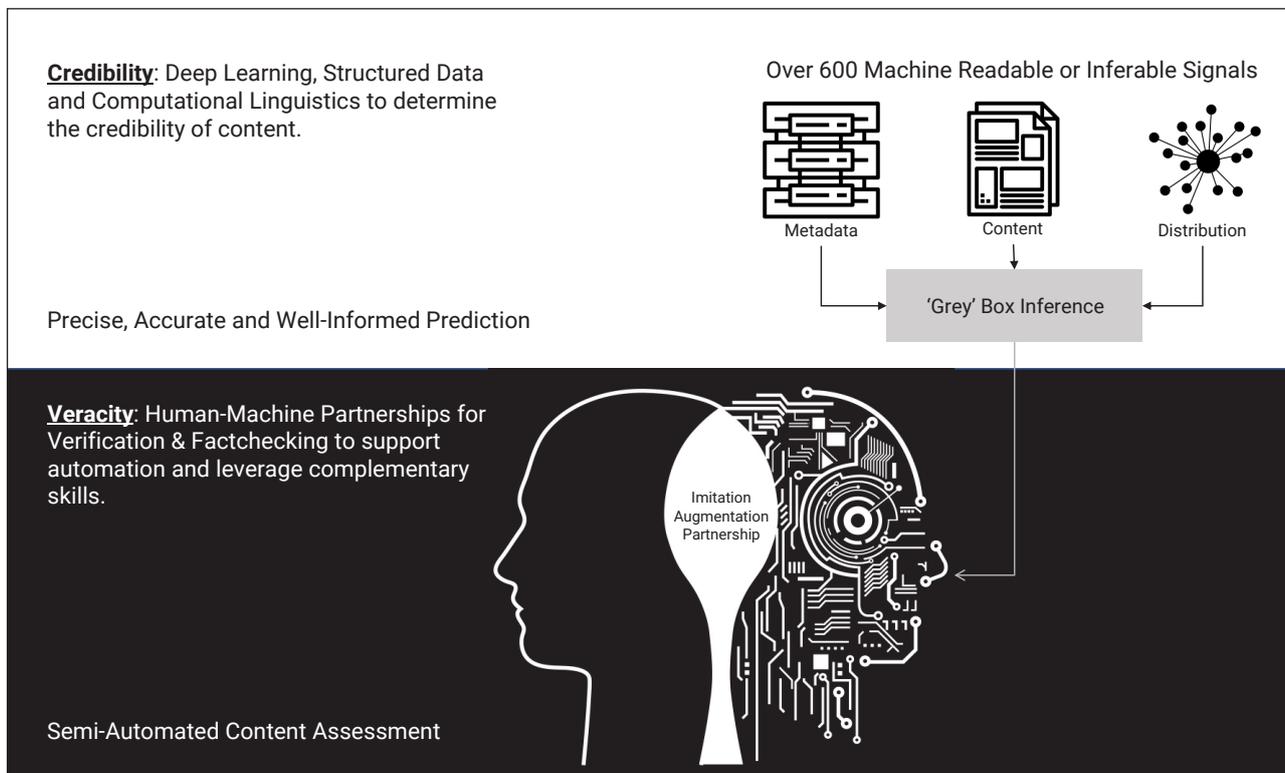
This is a vital component of our ambitions in the field, both in terms of our ability to stem the immediate flow of misinformation online,

and our long term ambitions to develop new technologies and approaches for a more sustainable and impactful fact-checking industry.

One of the biggest factors in the recent explosion in the volume and impact of misinformation has been its ability to propagate quickly online through social networks. Human fact-checking procedures are time-consuming and expensive, a huge obstacle both in terms of the impact and scale of fact-checking efforts. Whilst Logically's advances in the operational efficiency of human fact-checking processes are a positive step, full automation of fact-checking processes would enable a truly scalable solution to the problem of misinformation, and one which can negate the impact of misinformation by verifying, or debunking, claims in real-time.

We continue to develop machine learning and common-sense reasoning technologies that judge incoming new claims from our users against our universe of facts established by our fact-checking teams, single sources of truth and trust partners. This technology is already in action in the Logically app, where incoming claims follow a hybridised approach which first seeks to automatically verify a claim, relying only on human verification where this isn't possible. The development of this technology to enable full automation is one of our key priorities over the coming months and years, and a central component of our mission.

Credibility vs. Veracity



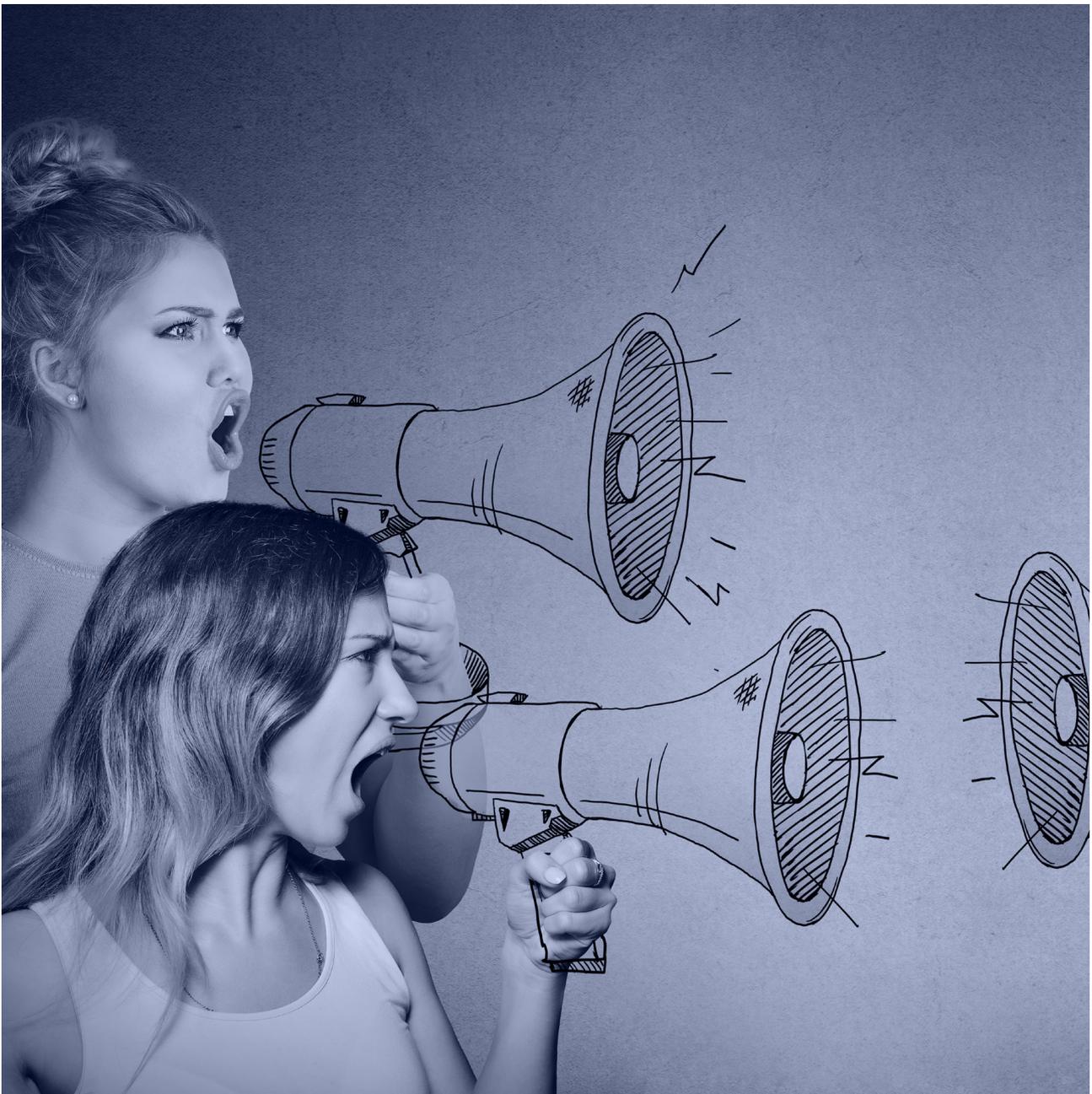
It's important to distinguish two of Logically's core capabilities: credibility assessment and veracity assessment. Our credibility assessment looks at indicators of content's overall accuracy, without specifically analysing claims contained within. We draw on content, distribution and metadata analysis to assess the content in its entirety. Veracity assessment, on the other hand, is concerned with a single statement or claim, conducting primary and secondary research (a "fact-check") to conclusively verify its factual accuracy. Put simply, credibility assessment represents an accurate well-informed guess approach to information verification, whilst veracity assessment enables an unequivocal assessment of a claim's accuracy.

Inevitably, the applications of the two capabilities are therefore quite different, with the most important difference - other than the judgement's certainty - being the speed with which the assessment can be carried out. Whilst Logically's credibility assessment is fully automated, veracity assessment requires human input and a critical mass of location or domain-specific facts to have been established in order to automatically evaluate a claim's accuracy, making a realtime judgement impossible at this stage.

Network Amplification

Understanding the interaction between multiple social networks such as 4chan, Reddit, Twitter and Facebook is crucial to understanding the dynamics of misinformation and disinformation. An internal research project which has yet to begin involves real-time monitoring of fringe networks such as 4chan, 8chan and Gab, and mainstream

networks such as Reddit, Twitter and Facebook to isolate the origination of claims and identify crossover points – when claims spread from niche groups and networks to broader communities and identify crossover agents – individuals/accounts responsible for pollinating major networks.



04

Credibility Assessment



**FAKE
NEWS**

**FAKE
NEWS**

Credibility Assessment

Logically uses natural language processing and understanding to process, understand and analyse text content and establish its credibility.

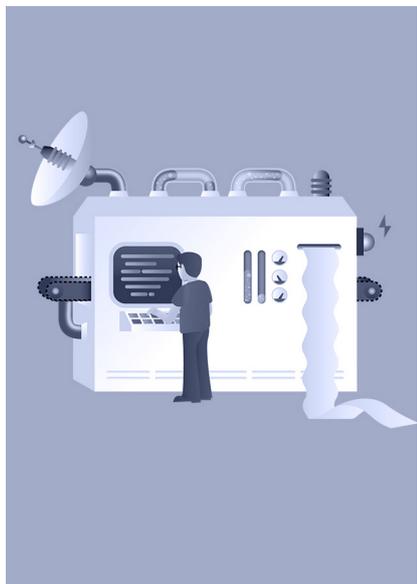
Over the last few years, a number of groups have emerged working to identify, define and categorise indicators of credible journalistic content. Some have taken highly academic approaches while others have done this work with a focus specifically on the applications that these indicators may have for technological solutions to misinformation and in an effort to return credibility and trust to journalism.

Groups such as The Credibility Coalition, The Trust Project and the Journalism Trust Initiative (headed by Reporters Without Borders) all aim to set new standards for journalists and technologists in the media. These groups look to pressure, not just journalists and publications into producing better, more accurate content, but also pressure digital platforms to prioritise more credible content on

feeds and search engines.

The emergence of these groups must be seen in the wider context of declining trust and widespread scepticism of traditional media. It is not surprising that more organisations are focusing on tackling these issues at the same time as others are looking towards the more specific problem of misinformation becoming pervasive online. The work of such groups also provides a valuable academic and professional framework on which to base emerging technologies aiming to alleviate these very serious problems. The Trust Project is of particular note as the standards that they promote for credible journalism have led to what they term “trust indicators”. By including machine-readable indicators of credible content on digital content it allows technologists to build tools capable of identifying instantly more reliable online content.

Several principles are central to Logically’s approach:



Extended Intelligence

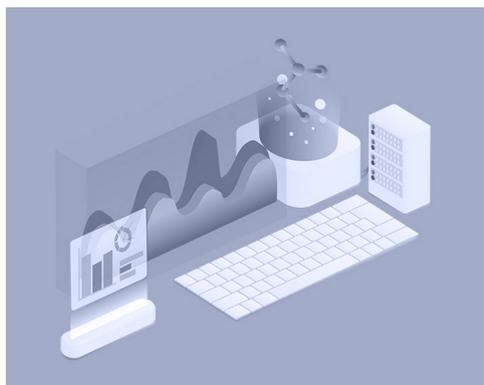
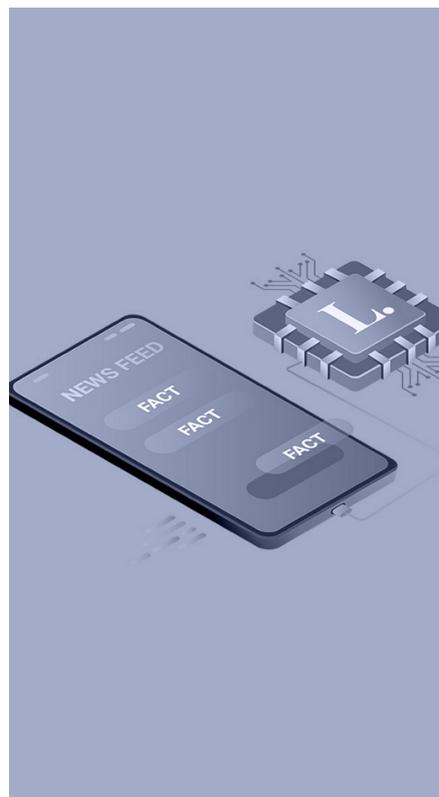
We believe that artificial intelligence should supplement human intelligence, not supplant it. Our technology works alongside our expert annotators and fact-checkers to evaluate news articles, and our products are designed to help our users consume information in a more discerning manner, not to abdicate them of the need to do so.

Furthermore “the truth” is often nuanced and sensitive, and empowering machines with absolute power to discern between truth and falsehoods would be irresponsible. An industry that has successfully deployed AI responsibly is Defence – using the human-in-the-loop framework wherein a person is responsible and accountable for pulling the trigger. We apply a similar methodology in our veracity assessment capabilities.

Explainable AI

The state of trust in AI at present can be summarised by the state of self-driving cars. If a police officer questioned a self-driving car or its driver for crossing a red light they would both reply with – “I don’t know”. This highlights many of the issues with the black-box nature of many AI applications today.

However, at Logically we believe for human and artificial intelligence to work together, they have to understand one another. This means doing away with ‘computer says no’-style intelligence, and developing AI capable of not only making decisions but justifying them too. Logically’s detailed annotation model and the hybridised analysis which populates it – analysing content, network and metadata to evaluate an article’s accuracy – means that any decision can be traced to its root justification. Our AI won’t simply brand an article as biased; it will identify the specific biases it detects to form this conclusion and where in the text they were found. Not only does this produce a more reliable and holistic conclusion, but it enables human verification of the artificial conclusion.



Hybridisation: A Three-Pronged Approach

Logically is unique in its combined analysis of network, content and metadata to reach its conclusions. Our AI leaves no stone unturned, evaluating every possible indicator of the overall article’s accuracy, as well as the specific claims contained within the text, to inform more sophisticated conclusions than rival models which rely on just one or two of these analytical channels.

Network Analysis

Propagation Analysis

Accurate, misleading and false information follow different propagation paths online after publication. We analyse the ways in which a story develops and proliferates on social media, looking for indicators which provide insights into the nature of the content itself.

Bot Detection

We can identify when a story is being spread by bots rather than humans based on common indicators we’ve identified. For example, a social media account that engages with and shares a story almost instantaneously is likely run by a bot, since human reaction speeds would be far slower. If a story is being spread by bots, then there is a high probability that its content is inaccurate and being maliciously propagated.

Nodal Analysis

We're building methods of detecting specific accounts or networks on social media websites that are more likely to propagate false or misleading content.

Metadata

We analyse the text's contextual surroundings online in order to gain insights relevant to an evaluation of the content itself. This analysis can be broken down into three categories:

Author level

Analysis of metadata surrounding a particular author and their biases, perspective, subject interests and track record of accuracy.

URL level

Analysis of metadata surrounding the webpage on which a piece of text is hosted. This will include analysis of that page's advertisements, tags and recommended content.

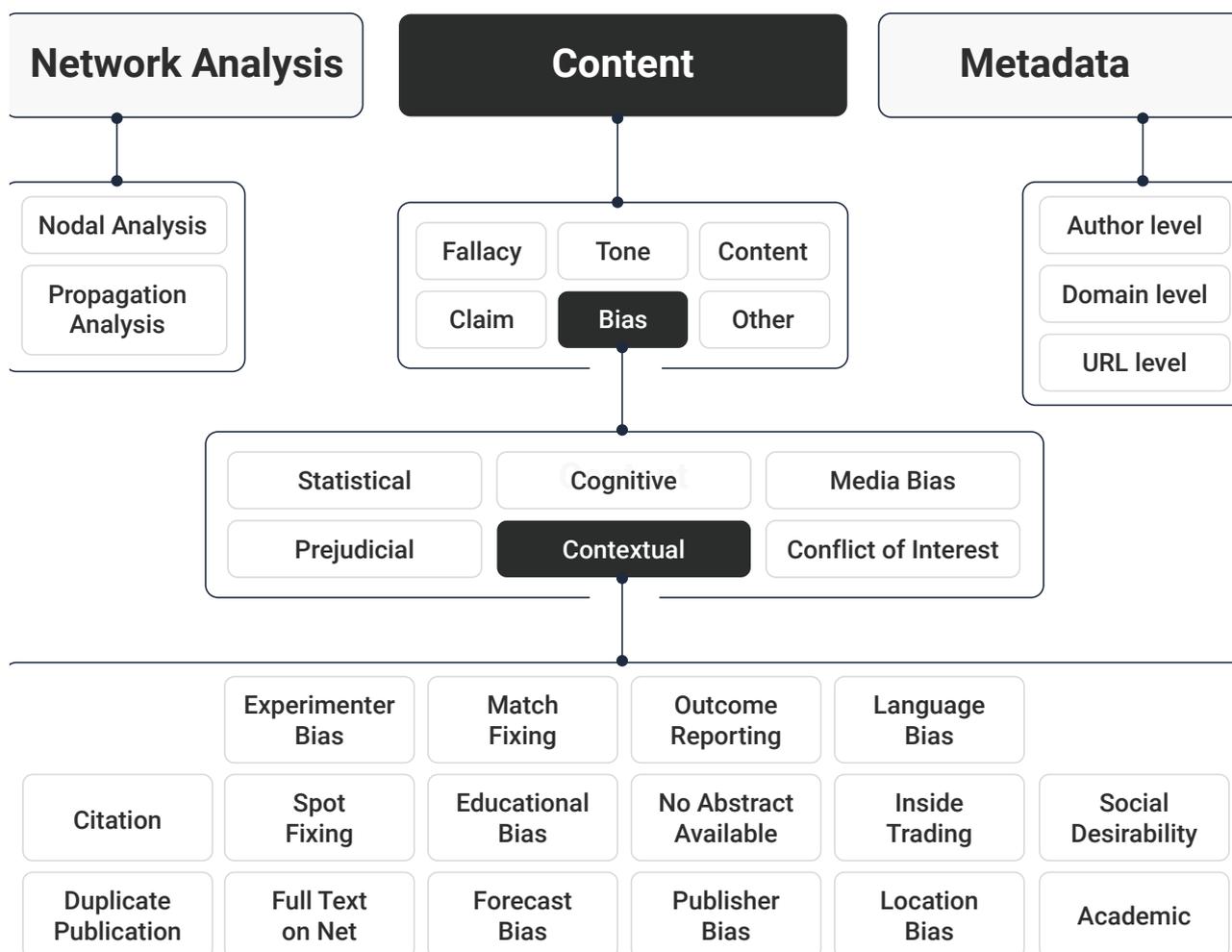
Domain level

Publisher-level analysis and domain health analysis. We can access spamblock lists and government databases relating to domain violations to evaluate the reliability of webpages hosted under that domain, as well as using our trust-linking capabilities to evaluate publishers themselves.

Content Analysis

Logically relies on various capabilities to analyse the content of a piece of information with great sophistication and detail. Our AI clusters indicators of misinformation under the headings of Fallacy, Context, Claim, Tone, Bias or Other, and can trace which indicators are identified within each article in order to achieve our goal of explainable AI. Under each category, there are various subcategories which outline even more clearly the precise flaw that is identified in the text. For example, we won't simply classify an article as biased, but rather by the type of bias exhibited: either statistical, cognitive, media-based, contextual, prejudicial or conflict of interest bias. As shown in the diagram (right), we further refine each of these subcategories; contextual bias alone contains sixteen sub-classifications. Our AI will quantify the level of each sub-classification contained within an article to inform its overall conclusion as to the article's veracity.

Explainable AI



Content Ingestion

Logically’s backend pipeline ingests content from over 100,000 sources, analysing 750,000 articles each week. These sources vary in perspective and format and are derived from a wide range of publishers, both traditional news outlets such as The Guardian, New York Times, Indian Express and Independent newspapers, and digital publishers such as The Huffington Post and Quint. Logically also monitors premium content from publishers such as the Financial Times through partnerships.

We take an impartial approach to content ingestion, eliminating human decision making which can be susceptible to bias from our ingestion process. We ingest content from all sources indiscriminately and without human moderation.

We additionally ingest shorter, unstructured content such as tweets and other posts on social media.

Credibility Assessment Methodology During Indian Elections

Logically deployed its credibility assessment capabilities for two months during the elections and assessments for each URL were logged both 1 minute and 1 hour after its publication. The automated assessments were validated by expert annotation and fact-checking teams. Over 100,000 websites from around the world were monitored for misinformation and other forms of problematic content during the month of April. Logically's credibility assessment capabilities were deployed across all content from these websites related to and/or originating within India made up 3,836 websites 174 of these were from mainstream media sources and 3,662 were from fringe outlets.

Findings During Indian Elections

From the period of 1st to 30th April 2019, Logically analyzed 944,486 English articles

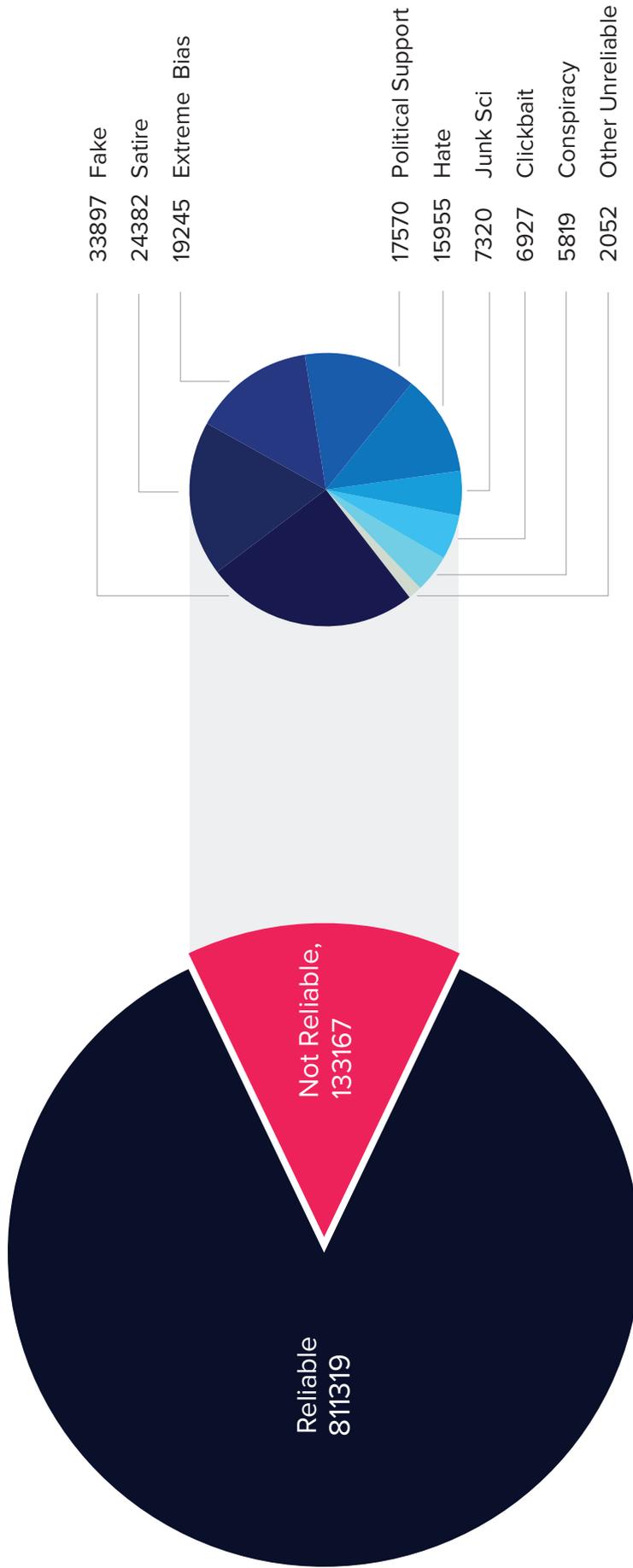
We found that 33,897 articles were completely fake. However the study also highlighted that 85% of English news being reported in the Indian media didn't contain any factual inaccuracies. Perhaps more significantly, the vast majority of the "fake" articles were from fringe websites and not mainstream media.

This might mean Indian media has been vindicated of charges of "fake news" and paid media, however if the problem of fake news is not taken seriously, then the incremental deterioration of trust in the fourth pillar of democracy will pose a serious threat to publishers and India's netizens. Furthermore, while mainstream publishers tended to be factually accurate the political biases in coverage were evident and in the cases of a handful of publishers were extreme.

The study pointed that during the same period, fake news pieces were shared more than 1 lakh times, hateful articles were shared more than 3 Lakh times and 15 Lakh shares were connected to extremely biased stories likely reflecting the sharer's personal opinions on topics. As a result, readers could be entering filter bubbles and echo chambers on their own.

Hateful and Junk Science articles were shared at significantly higher rates on Facebook (20 and 34 shares per article) and 'fake' articles were shared at significantly lower rates (3 per article) suggesting the platforms' efforts at downranking misinformation and already fact-checked content have been very successful. This further highlights the unique traits of mis/disinformation in India – the most significant platforms for problematic content have been closed networks such as WhatsApp, Private Facebook Groups – more on this in the Veracity section.

No. of Articles Published in April by Class



Shares on Facebook per article by content type



Credibility Assessment: Evaluation

We constantly evaluate and iteratively improve our models, which are tested according to the yield they attain. We benchmark automation performance against that of our expert annotation team, producing an F score which reflects both the recall and precision of our models. We've been delighted with the performance of our models in each of the three areas of analysis we conduct:

- Content Analysis: 92%
- Distribution Analysis: 93%
- Metadata Analysis: 88%

When ensembled these deliver a level of performance significantly beyond the current state of the art approaches – over 95%.

Credibility Assessment: Use Cases

Logically's Credibility Assessment provides a well-informed prediction of content credibility within seconds. In sensitive and high volume applications such as election security and platform content moderation, these can be used as a 'funnel' for veracity assessment that also provides an initial heuristic to aid assessment.



TECH • CYBERSECURITY

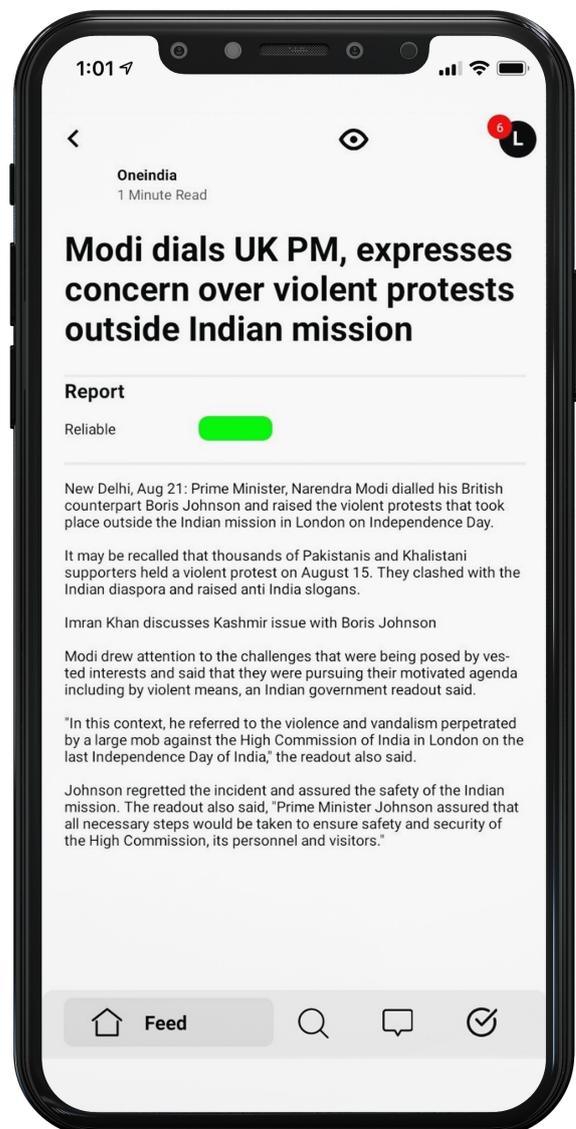
Hoax Over 'Dead' Ethereum Founder Spurs \$4 Billion Wipe Out

Algorithmic traders and select Hedge funds trade either exclusively or primarily based on online and social data. However as has already been established, the signals and event data received through these channels can be unreliable or worse - manipulated by adversarial traders. A low latency application of the same credibility assessment capabilities extended to cover finance in addition to current affairs can be deployed as a risk mitigation mechanism for algorithmic traders.

The capabilities as they stand, offer an affordable and fully automated solution for less sensitive but high volume use cases such as inventory moderation for ad-networks and brands. These can also be extended to cover 'fake reviews' for marketplaces and review platforms.

Logically's flagship product, a news platform which automates the intelligent curation, generation, contextualisation and verification of news within a direct-to-user mobile application, was first launched in India in April 2019, and then in the UK in August. Our credibility assessment capabilities are a vital component of this application, informing several key features:

- Logically's automated fact-checking capabilities are informed by our credibility rating for the source as an initial guide to the fact-check-worthiness of a specific claim.
- Logically is a story-centred news feed; we organise the news according to our story clusters which aggregate all the content we find relating to a story. Within each story, we suggest individual articles for users looking to consume deeper or opinion-based content on a story; each article is analysed by our technology, with any indicators of an article's credibility (or lack thereof) highlighted for the user to help them consume the news in a discerning manner.
- Logically's credibility ratings inform our objective story summaries, helping our technology pick out the facts associated with a story by prioritising the articles from which these facts can be taken.
- Articles within the Logically news experience are also labelled and users are warned before seeing any articles with problematic traits.



05

Veracity Assessment



Methodology: The Life Cycle of a Claim



Logically's boasts the world's largest team of dedicated fact-checkers, supported by in-house journalists, innovative technology and efficient, streamlined processes designed to safeguard the integrity of our fact-checks whilst maximising their efficiency. While we aspire to a fully automated fact-checking solution, we have developed an hybridised process which supports the development of our fact-checking algorithms, enables the incremental adoption of our automated fact-checking technology as it matures, and enables efficient and high-quality fact-checking in the meantime.

1. Incoming Claim

Users can submit suspect claims for fact-checking that they encounter within the Logically app, or from third party publishers by sharing the article with Logically or pasting the URL into the app. Once we have the article, our claim detection technology gets to work extracting the factual claims within, enabling the user to select the one they'd like verified and submit it to our fact-checkers. Users can also enter raw text or paste a message/post from another platform.

Claims then appear on the bespoke dashboard developed to help our team progress claims through our process efficiently, maximising our chances of preventing false claims from spreading.

2. Automated Factcheck

Incoming claims are first checked against our single source of truth databases through a vectorised analysis of the incoming claim against related claims. This represents an initial processing stage of a larger scalable automated fact-checking solution which will become more effective as our universe of known and inferred 'facts' is expanded through human fact-checking and sourcing from single sources of truth databases such as Government datasets.

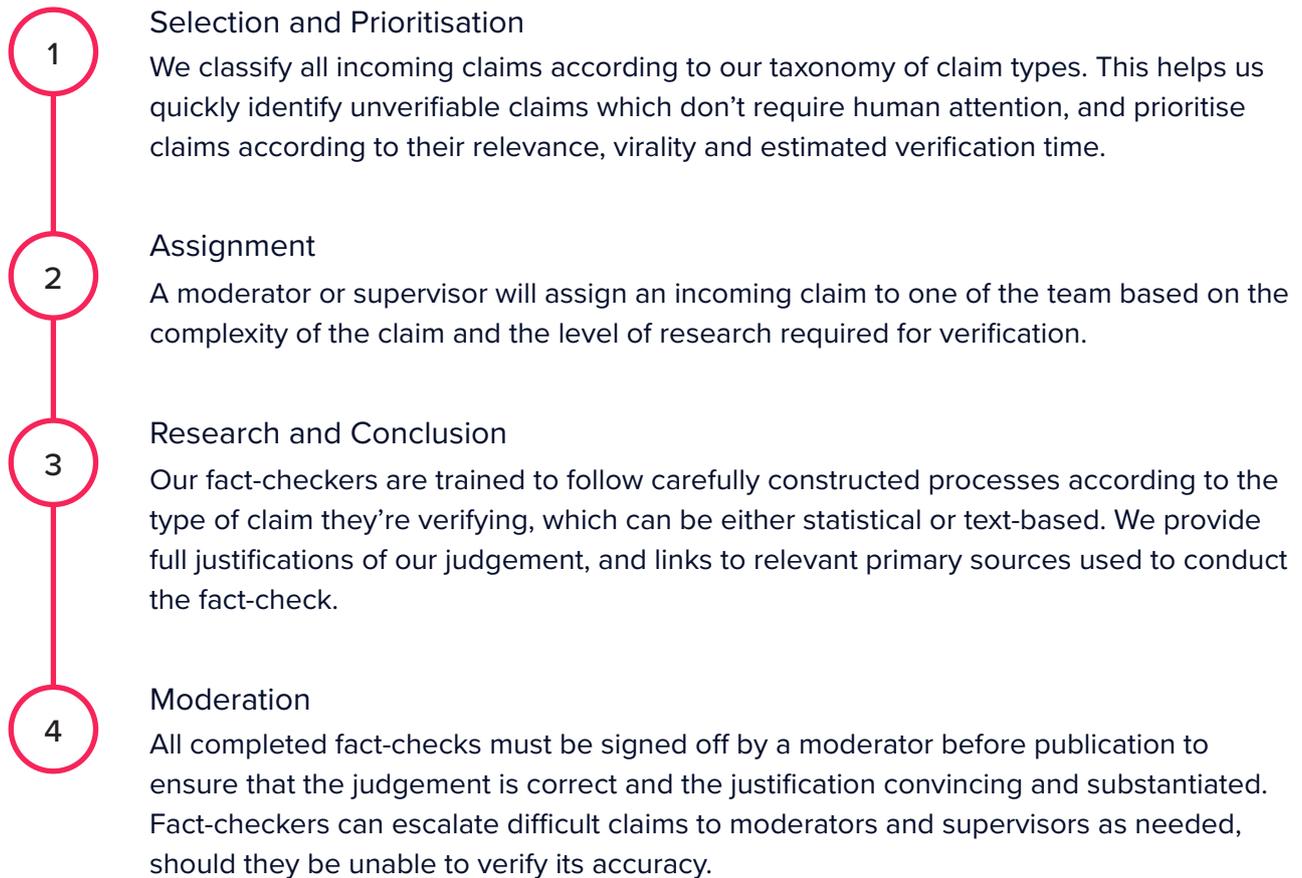
Additionally, users are able to determine if images are manipulated via the Logically app. The analysis of the image is returned to the user within seconds with colour coded highlights on all areas the models believe likely to be manipulated. Users can forward the same image for fact-checking if they would like the content of the image itself verified.

3. Random Sample Verification

Logically conducts random verifications of claims fact-checked through our automation capabilities to ensure that any errors are detected and used to develop the algorithms which conduct the fact-check.

4. Human Fact-check

All claims sent to our human fact-checking team are subject to the following process:



5. Judgement

Finally, we publish our verdict, which will be either: True, Partially True, False, Misleading or Unverifiable. Logically's verdicts have been carefully designed to encourage sharing and maximise virality; the judgement and claim are featured on an eye-catching image designed to look great on social media platforms. Through these viral verdicts, we hope to maximise the impact of each fact-check and help ensure that any false claims are debunked as broadly as possible.

Methodology During Indian Elections

Logically's verification work during the 2019 Lok Sabha was conducted by our 27 person team of fact-checkers, moderators and supervisors. The team worked full time in order to check as many requests as possible, putting in countless hours of overtime to make sure that we could verify as much as possible.

In keeping with the mission of Logically, in helping users to access better quality information and engage in constructive, empathetic debates around current affairs, our fact-checking operation prioritises user requests for verification. This means that the majority of claims that we check have been submitted by individuals through the Logically app. Our fact-checking and editorial teams will also sometimes find claims that they'd like to check, if they feel that the issue is particularly important or lacks adequate coverage by other organisations.

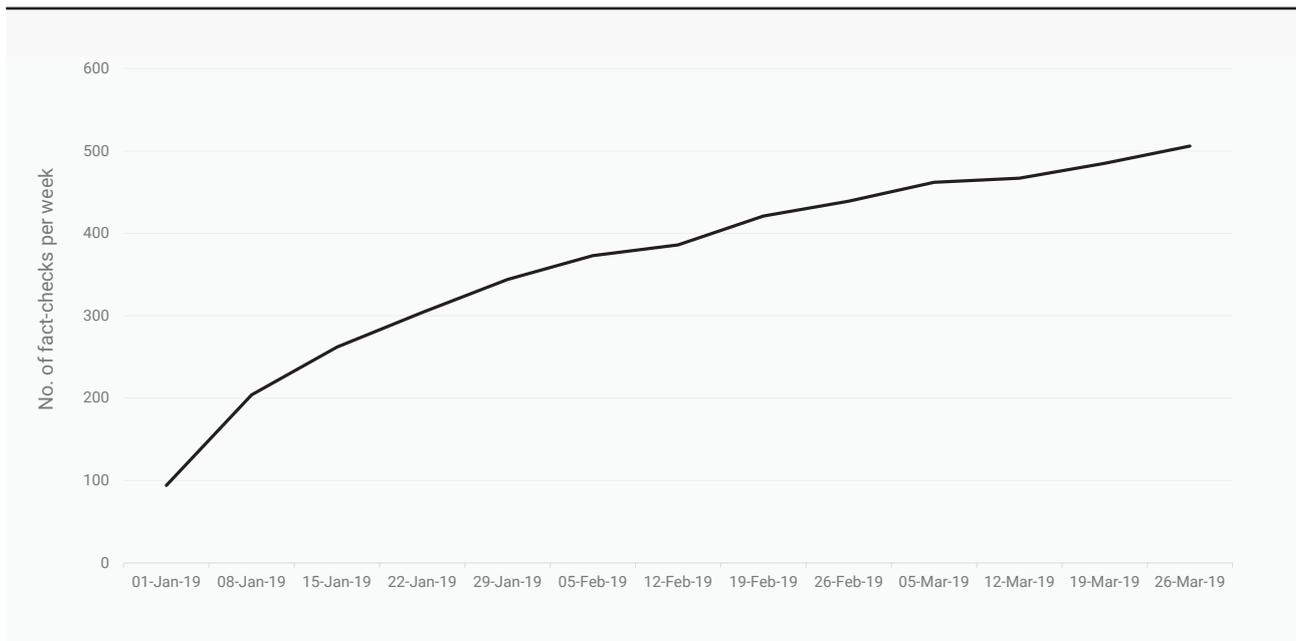
Once the claim has been researched and our team reaches a verdict, a Safe to Share image is created. This has been designed to include an image, easy-to-read verdict, quick summary of the fact-check and then a longer explanation of how it was conducted. This format was inspired by the success of Verificado, a pop-up fact-checking operation during the Mexican general election in 2018, which adopted innovative storytelling techniques in order to optimise its work for social media audiences. Our own Safe to Share images have been developed specifically with WhatsApp in mind, reflecting the critical role that the platform plays within the Indian information ecosystem and its recent struggles with "fake news."

Findings During Indian Elections

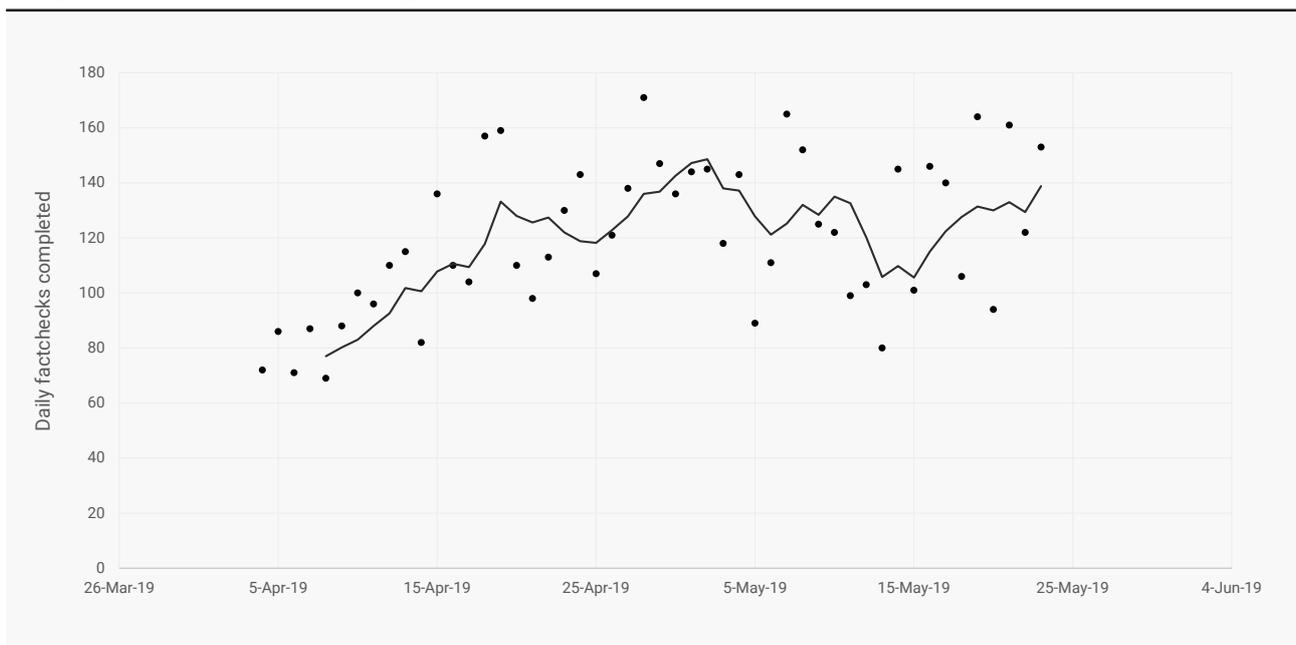
Our team of fact-checkers has been trained using industry-leading standards and methodology from organisations including Google, Poynter, First Draft News and the International Fact Checking Network.

With turnaround time a key performance metric, given the fast-paced nature of the news cycles, our 27-person team of fact-checkers have continued to improve in efficiency since the commencement of the project in January 2019. After an initial month of training, the team to date have completed approximately 6,000 fact-checks on everything from current affairs to sports and celebrities, defence, economics, politics and science. Each member of our team has gone from producing approximately one final fact-check a day to producing six. Complete oversight from our moderators, as well as a rigorous system of checks and balances and emphasis on excellent sourcing of evidence ensures that we produce high-quality fact-checks as quickly as possible, to stop misinformation before it spreads.

Weekly Factchecking Output - Pre Election

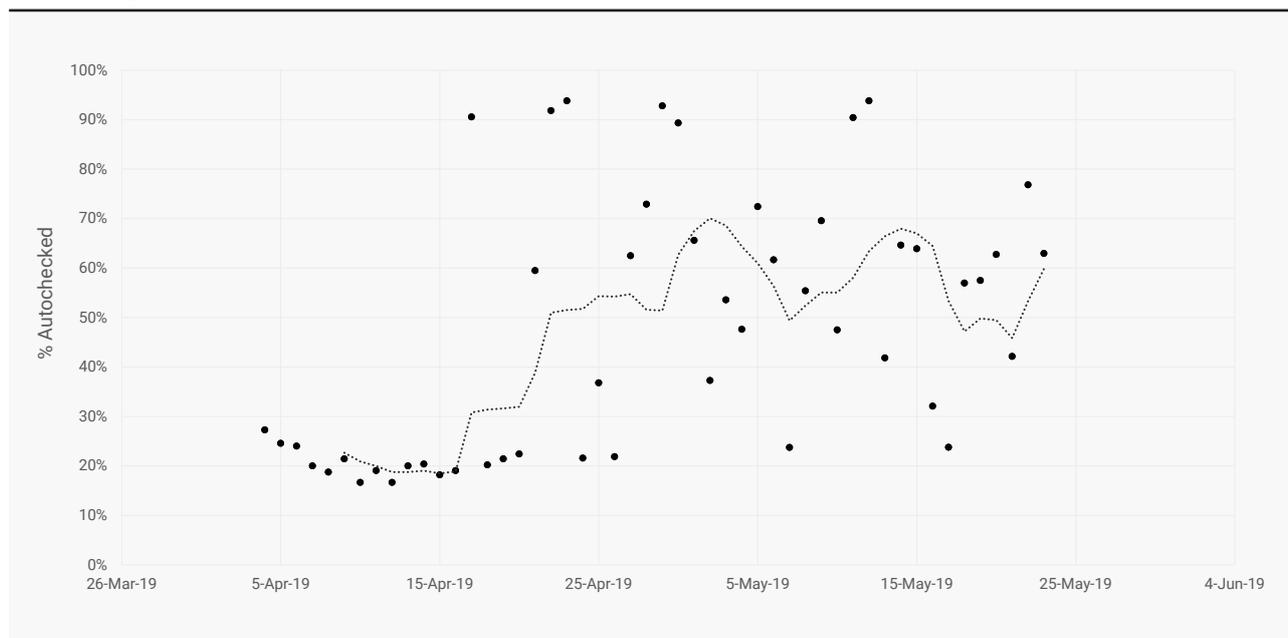


Daily Human Factchecks During Elections



During the initial months before the election the improvements in productivity from updated dashboards and tools take time to materialise - following a typical learning curve that appears to stagnate at the 120 per day level once the election is in full swing. This level of output suggests maximum feasible productivity given the current state of supporting technologies has been achieved.

Percentage of User Requests Responded to with Autochecks



During the initial weeks of the deployment, the automation capabilities answered just 20% of incoming claims, however they significantly improved over time and in the month of May once a substantial database of facts was established and a few optimisations were made to the automation systems, they dealt with 40-70% of all incoming requests.

Perhaps more importantly, the automation systems conducted 90% of fact-checks on high traffic days. These occurrences highlighted by the chart occur on days where over 500 geographically and topically concentrated requests were received.

Evaluation

The objectives of our fact-checking operation in India were ambitious, with a very short turn-around time between implementation and the commencement of the 2019 Lok Sabha elections. A user-centric model of verification, in which we prioritised user requests for clarification or fact-checks, meant that our team was reliant on the Logically app at the same time as it was being launched.

At particular points during the project, these factors acted as temporary constraints on our fact-checking output. A tight timeline posed some challenges as we assembled and trained the team, meaning that there was less room for error and more pressure to learn and adapt to a challenging role quickly. Technical issues with the app in its beta phase also imposed some structural limitations upon the team's operations.

However given the demanding nature of the project, the team rose quickly to these challenges, rapidly adjusting to political developments in India and incorporating these insights into their work. The result was an acceleration in fact-checking output to approximately 600 fact-checks per month in the first three months.

As we begin to expand our fact-checking operations to include claims from the United Kingdom, this rate continues to improve, and our team of fact-checkers, with the support of complete editorial oversight from Logically London, is quickly developing a thorough understanding of the UK political and information environment. We hope that these skills can be implemented effectively in time for the 2020 US presidential elections, as well as any impromptu elections in the UK.

06

Future in India



Supporting the Fact-Checking Ecosystem

Logically will use the technological and operational expertise it gained in launching its consumer application and third-party fact-checking partnerships during the 2019 election as a launchpad in the fight against misinformation in India.

Logically boasts the world's largest team of dedicated fact-checkers, with around 27 fact-checkers focused on accurately and efficiently verifying claims sources through the Logically app. This, and third party partnerships with Election Commissions, Police, Social Media Platforms and Messaging Platforms, underpins our strategy in the field.

The accuracy and integrity of our fact-checking is our biggest priority. We've established rigorous processes for the assignment, research, justification and approval of our fact-checks, overseen by our UK-based editorial team familiar with the standards imposed by the IFCN. Our fact-checking team have been trained in line with these standards, and our own expectations

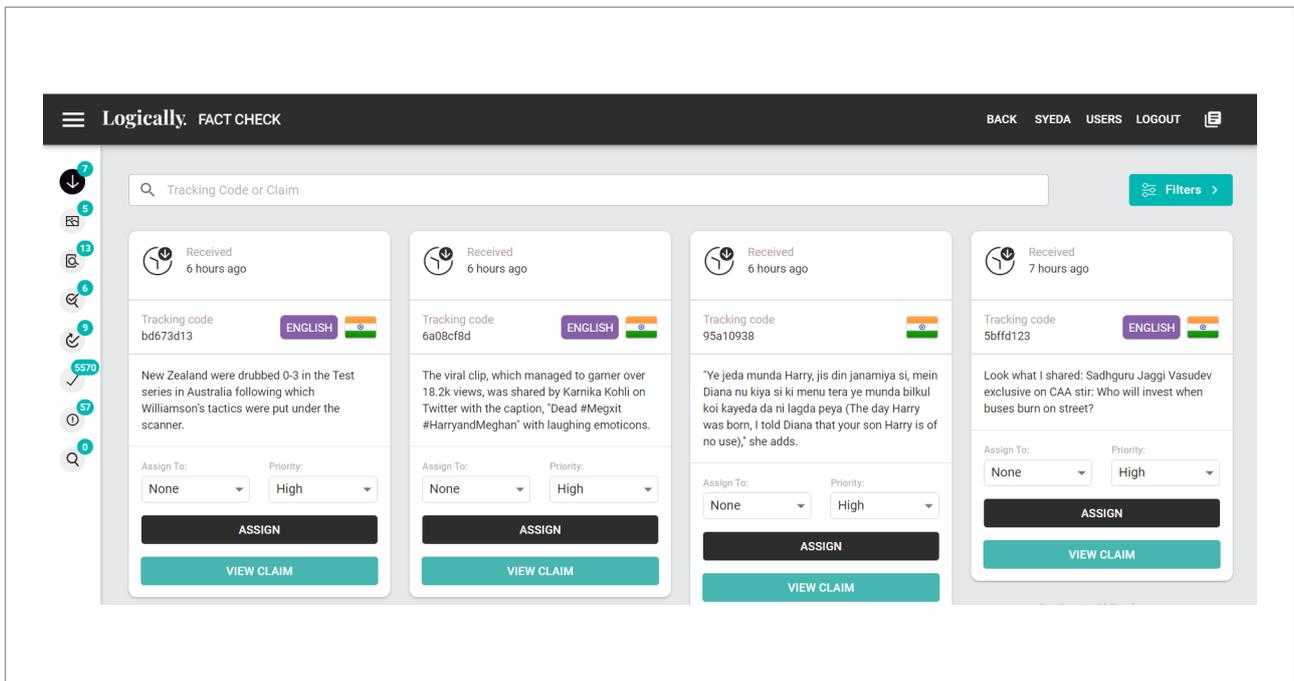
Industrial Revolution in Factchecking

Logically will deploy its expertise to facilitate an overdue technological revolution in the fact-checking industry. This is vital to our overall attempt to stem the growing misinformation problem in India; if the industry is to keep up with the speed at which misinformation propagates online, innovative means of improving the efficiency and reach of fact-checks will be essential. We aim to bring technological innovation in three main areas:

Efficiency

One of the difficulties of the current fact-checking industry is that conducting a fact-check to a high standard is a time consuming process. In a sharing nation where information can multiply and propagate almost instantly, it is almost impossible to negate the effects of the false article by the time it has been rigorously fact-checked.

Logically is developing various technological solutions to enhance the efficiency of fact-checking operations, including its fact-checking dashboard. This dashboard provides an end-to-end solution which helps fact-checking organisations manage the entire process from assignment to publication. Already deployed in-house, this dashboard will eventually be rolled out as an off-the-shelf solution for third party fact-checking organisations. The primary objective of this product is not the immediate generation of revenue, but access to a wider variety and volume of completed fact-checks to support the next component of Logically's technological revolution: automation.



Automation

While efficiency tools such as our fact-checking dashboard will be a useful step in helping the fact-checking industry combat misinformation, automation of these processes is the elusive goal which would truly enable it to solve this issue in a proactive, rather than reactive way, quashing false information before it spreads rather than facing an uphill battle to debunk widely believed falsehoods.

Logically has already made significant progress in this area, developing machine learning technology capable of matching new claims to be fact-checked according to previously established facts. This technology is at an early-stage and not yet ready to offer a sustainable and independent alternative to human fact-checking, but its continued incremental development is dependent on the provision of human fact-checks on which its algorithms can be trained, and which expand the database of known facts from which it can verify future claims. Our engineering team have also developed proprietary algorithms which review incoming claims, seek to verify them automatically, and send fact-checks which can't be automated to our human fact-checking dashboard.

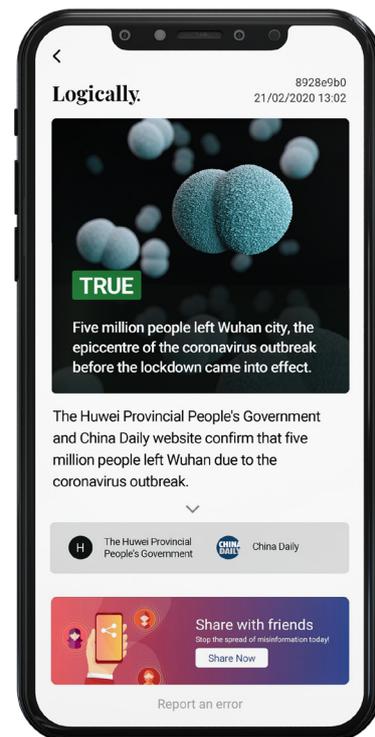
Logically's human fact-checks, combined with off-the-shelf products which should enable access to a greater volume of fact-checks, are vital to the development of this technology, which will gradually become a more important component of our overall fact-checking efforts.

Virality

One of Logically's priorities in establishing its fact-checking procedures was the format of fact-checking outputs. We believe that in order to combat misinformation, the fact-checking industry must harness the same virality which makes misinformation so successful in digital arenas.

We automatically render our fact-checks as shareable graphics designed to maximise the virality of the output and encourage its spread on social media. We hope that this can improve adoption of fact-checking services to identify false claims more often and faster, and also improve the spread of that specific judgement to minimise the misleading effect of the initial misinformation.

Through these three areas of innovation, Logically expects to carve out a prominent role in the broader fact-checking industry and fight against misinformation in India, leading the way in bringing innovation to the industry and facilitating a more sophisticated solution to a growing problem.



Leverage Capabilities During Elections

Election periods provide unique opportunities and challenges for Logically. Not only does societal and platform interest in the issue of misinformation increase drastically, but the volume, virality and impact of misinformation explodes alongside, with malicious actors doubly incentivised to publish false claims in pursuit of political or ideological gain.

Democratic processes can be particularly susceptible to manipulation during election periods. The high political and ideological stakes can attract manipulation and interference from actors within a democracy and malicious foreign parties.

The impact of such interference can be extremely damaging, both diminishing trust in the news and reducing the legitimacy of electoral mandates. Famously, the Brexit and US elections in 2016 exemplify these dangers; accusations of misleading claims during the referendum campaign and Russian-led disinformation campaigns around the US election have undermined the legitimacy of both elections.

Propagation of misinformation and other forms of harmful digital content is similarly amplified during election periods; in a digital world where engagement, best generated through sensationalism, is the primary determinant in the propagation path of a piece of content, elections and the corresponding surge in political interest create the perfect conditions for the spread of misinformation.

Logically's Immediate Objectives in India are to:

Support Democracy

Logically will be proactive in supporting democratic processes during these periods, investing resources to ensure that our unique capabilities are deployed effectively during elections. We will enter strategic partnerships with relevant governmental bodies, digital platforms and other organisations to help debunk misinformation through our fact-checking capabilities, and deploy our viral monitoring and network analysis capabilities to contain its spread. We consider this increased investment and strategic attention during the pressure points that elections generate in the information ecosystem to be a vital and indispensable component of our overall mission.

Build Awareness of Logically Capabilities, Products and Services

Whilst immediately challenging in terms of Logically's mission, we believe we can leverage election-oriented interest to build brand awareness by taking a prominent and successful role in fixing problems relating to misinformation during these periods. Through this role, we will establish our brand as a mainstay of India's news industry, both in terms of our news platform and our fact-checking activities.

Build Platform, Government and Media Partnerships

Platforms are increasingly expected to be proactive in preventing the propagation of misinformation through their platforms and have recently experienced the potential reputational damage of failing to fulfil this expectation. Logically's partnerships with Platforms, Election Commissions and Newsrooms, verifying claims are a potent example of collaborative efforts to fight mis/disinformation while respecting the priorities of each group of stakeholders. We look forward to being more ambitious in the near future and build valuable partnerships which will sustain our commercial and social objectives beyond election cycles.



Live Fact-checking for Broadcasters

We have spoken already about methods of delivering fact-checks and verified content back to audiences online and on messaging services - but how do we do the same when trying to verify claims made via broadcast television or live-streamed content? In these instances, latency is a huge issue. If claims seen or heard live go unchecked then it can be very difficult to deliver corrections to the audience who received misinformation.

It is therefore prudent to consider how we can attempt to fact-check live content in real time. Broadcasters themselves have little internal capacity to verify and fact-check information at scale. This is especially true in today's 24-hour news cycle and general declining public trust in the media.

Logically's approach to verification can help ease pressure on broadcasters and also keep a check on live-streamed content. It is arguably the most important to provide high-quality and verified content at times of intense political and democratic flux. The intense balloting and campaigning during election periods for example sees an increase in the importance of and a number of televised debates, panels and rallies. By using live fact-checking techniques, experts and politicians can be held to account during the times it matters the most and in doing so also provides broadcasters with new, innovative programming to support editorial their editorial offering and simultaneously re-establish credibility and trust in their brand.

Live Fact-check Speeches by Politicians during Interviews and Rallies



Live Fact-check Debate Show Panellists



07

**WhatsApp
Solution for a
Sharing Nation**



SHARE

Logically's solution to the problem of misinformation in India seeks to harness the rapid digitalisation of India's information ecosystem; India has become a sharing nation, an interconnected web of information disseminators pursuing individual goals ranging from humour to ideological conversion. These conditions have greatly exacerbated the problem of misinformation, and whether we like it or not, they're the conditions in which we must find a solution to that same problem.

India has undergone something of a digital revolution in recent years, with consumer technology growing at a remarkable rate. Smartphone ownership has risen from 33% to 40% since 2017^{22,23}, reflecting a broader and even more significant explosion in remote data usage. In 2017, 108 Petabytes of data was transmitted online. This figure is expected to grow by 490% by 2022, with the nation consuming 646 Petabytes.²⁴

Unsurprisingly, social media companies have benefited from the nation's growing connectivity, with the nation's active social media users growing by 24% from 2018 to 2019, reflecting a consistent trend which is expected to continue in the coming years.²⁵

Such growth in misinformation-friendly platforms has had inevitable consequences on the information ecosystem in India. WhatsApp groups have emerged around ideologically homogeneous groups dedicated to sharing information which supports their worldview. These groups, often with hundreds of members, become echo chambers congenial to the spread of misinformation; an ideologically-blinkered network of potential propagators ready to embrace content

supporting their common worldview.

The role of social media in spreading misinformation is well established; the engagement-oriented content discovery algorithms incentivise sensationalism, oversimplification and gossip rather than detailed and balanced journalism. This was prevalent in the Indian election, which featured a proportion of polarizing political news topped only by the US election in 2016. More than a quarter of the content shared by the election's victorious party, the Bharatiya Janata Party (BJP), was "junk news" of some form.²⁶

The combined efforts of the academic, technology and media industries have yet to find a solution to misinformation capable of functioning in this environment. Attempts have thus far had one thing in common though: they fought, rather than embraced, the nation's orientation towards shareable online content. Whether encouraging user verification and responsible dissemination of information or imposing those responsibilities on the platforms themselves, solutions have always been reactive, seeking to counteract the virality phenomenon.

Logically has flipped this approach on its head, creating a simple yet innovative solution to misinformation which, based on early data collected during the 2019 Indian election, is highly effective. Our role in combating misinformation is engineered to harness the same benefits of the modern information ecosystem which enables misinformation's spread in the first place, achieving high virality and engagement through the format and strategic publishing of our fact-checking output. We produce shareable graphics

22. 'Digital 2019: India', We are Social and Hootsuite. 31st January 2019.

23. 'Digital 2017: India', We are Social and Hootsuite. 1st February 2017.

24. <https://qz.com/india/1483368/indias-smartphone-internet-usage-will-surge-by-2022-cisco-says/>

25. 'Digital 2019: India', We are Social and Hootsuite. 31st January 2019.

26. Vidya Narayanan, Bence Kollanyi, Ruchi Hajela, Ankita Barthwal, Nahema Marchal, Phillip N. Howard; 'News and Information over Facebook and WhatsApp during the Indian Election Campaign', Data Memo 2019.2, Oxford, UK: Project on Computational Propaganda.

featuring a headline image, attention-grabbing judgement, and a concise justification, providing all the details required by the IFCN's fact-checking regulations.

We've also entered into strategic partnerships which will maximise the exposure and virality of our fact-checks, extending their reach beyond in-app verification for Logically users. Deployment on third-party apps (most notably in the context of India, WhatsApp) provides a means of accessing Logically's verification capabilities through immensely popular platforms and will enable us to upscale our fact-checking operations and underpin an even more ambitious solution to misinformation in

India. Moreover, we'll be able to access the ideologically-affiliated groups which play such an important role in the national discourse, debunking claims and stifling the radicalisation that such polarized and homogenous communities could foster.

Logically is committed to fighting the issue of misinformation in India going forward, and data collected around the 2019 election supports its theory that India is and will remain a sharing nation, and organisations dedicated to addressing the growing misinformation crisis must find solutions within this environment.

Long-Standing Questions

As suggested previously, WhatsApp is likely to be the most significant avenue by which misinformation is circulated in India. Significant not just because of the frequency in which it is used but also because of its encrypted nature, also previously touched upon. These two features of WhatsApp usage in India also pose two distinct, yet closely linked challenges:

1. How can we get access to the kind of information being shared in order to spot problematic content, whilst respecting the privacy of users and without compromising encryption?
2. How can we find a way of disseminating credible information, verifications and debunks, back to the same audiences who are receiving this content in the first place?



Related Work

Previous work has been done in attempts to address misinformation during elections, specifically with messaging platforms such as WhatsApp. Arguably the most successful of these projects took place during the elections in Mexico in 2017, which turned into an unprecedented fact-checking collaboration. In all the collaboration involved over 100 journalists representing about 60 different media partners. The project, named Verificado – Spanish for ‘Verify’ – looked into ways that they could most effectively spread fact-checks and debunks online, and how they could ensure that this content reached those most affected by misinformation.[1]

One aspect of Verificado’s project was to set up a public WhatsApp account to which members of the public could send requests for verification of individual pieces of content. Once received by the Verificado team the content was verified and graphics were made up to illustrate the team’s verdict of that claim or piece of content, along with the official Verificado logo and a timestamp. This was then published on various platforms and sent to other organisations involved in the collaboration who in turn published it on their own platforms.

By the end of the election in Mexico, the Verificado brand had amassed a widespread following amongst voters. The official website had received over five million views and the team had published around 400 fact-checking notes and around 50 videos. These videos were also posted on numerous platforms and websites, some receiving more than a million views each. Most importantly when we consider transferring similar methods of combatting misinformation to an Indian context, the WhatsApp group had garnered more than 10,000 contacts.

Exposing Misinformation on closed networks

Individuals on closed networks such as WhatsApp are exposed to disinformation as a result of organised propaganda campaigns and inaccurate viral stories. As these networks tend to be encrypted, a way to obtain access to content while respecting platform policy and user privacy would be to encourage whistleblowing on these networks. By encouraging the individuals themselves to share content they have received, fact-checkers and other organisations gain instant access to problematic content which would otherwise go unchecked. The question therefore is; how do we motivate this sort of whistleblowing?

There are a number of obvious motivations for individuals to share content from their closed network. Concern for their communities and networks, sparked by seeing content that they are

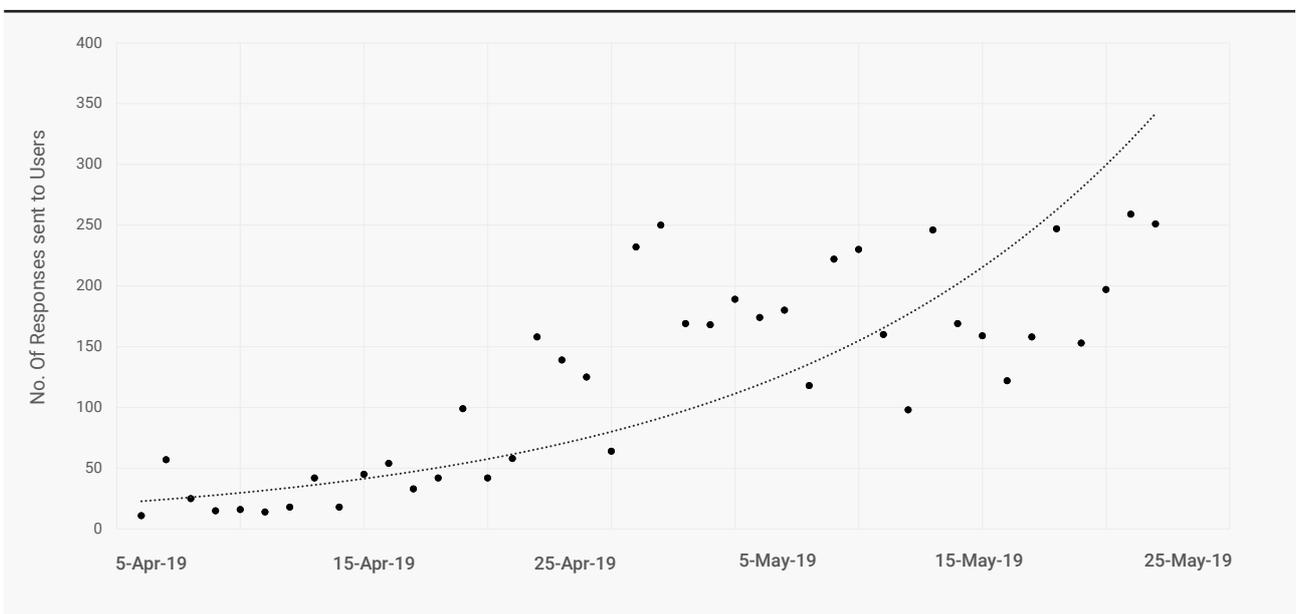
1. <https://verificado.mx/>

suspicious of is probably most likely to instigate a whistleblowing response. Either this or out of a sense of activism spurred on by seeing negative information targeted at.

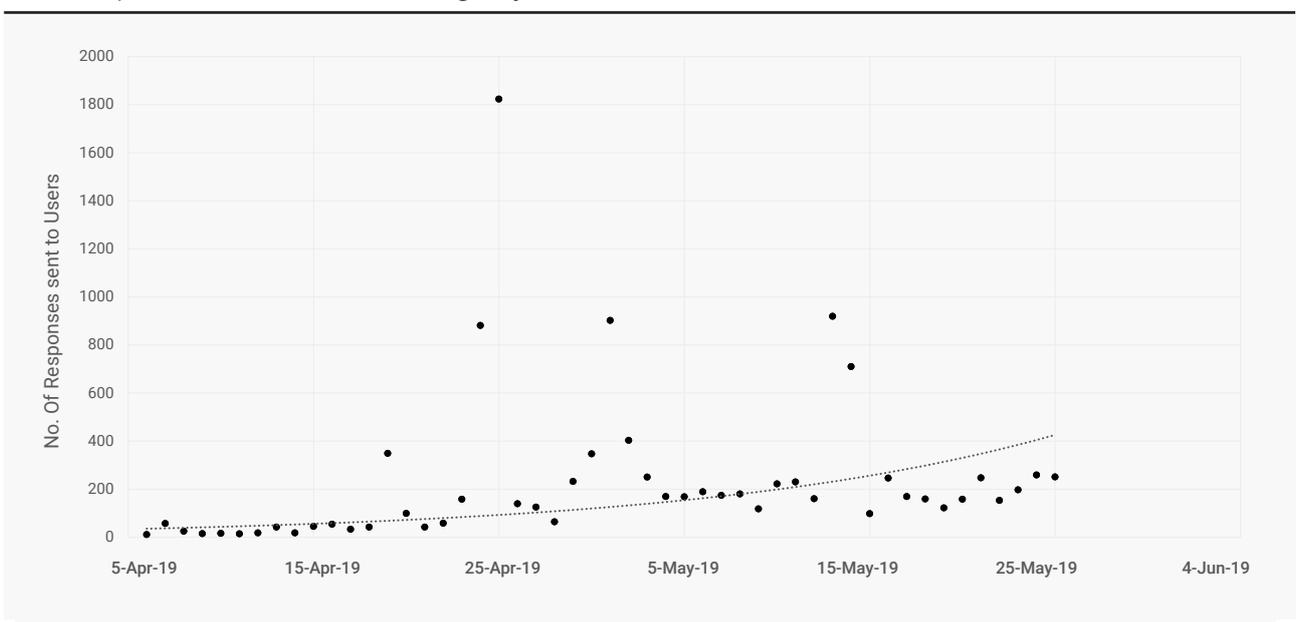
These motivations are also broad enough within news consuming audiences as the primary psychologically gratifying motivations for news consumption are information foraging, socialising and status-seeking.

By establishing a network of motivated and active “forwarders” who are consistently exposed to mis/disinformation, we can gain access to highly shared questionable content. In the 50 days between 4th April 2019 and 23rd May 2019 11560 unique user requests were received. The level of activity during non-key dates [key dates are days of and the day before an election stage] shows an exponential rise over the 50 day period implying user satisfaction and improving operational capacity.

User Requests Answered - Without Key Dates



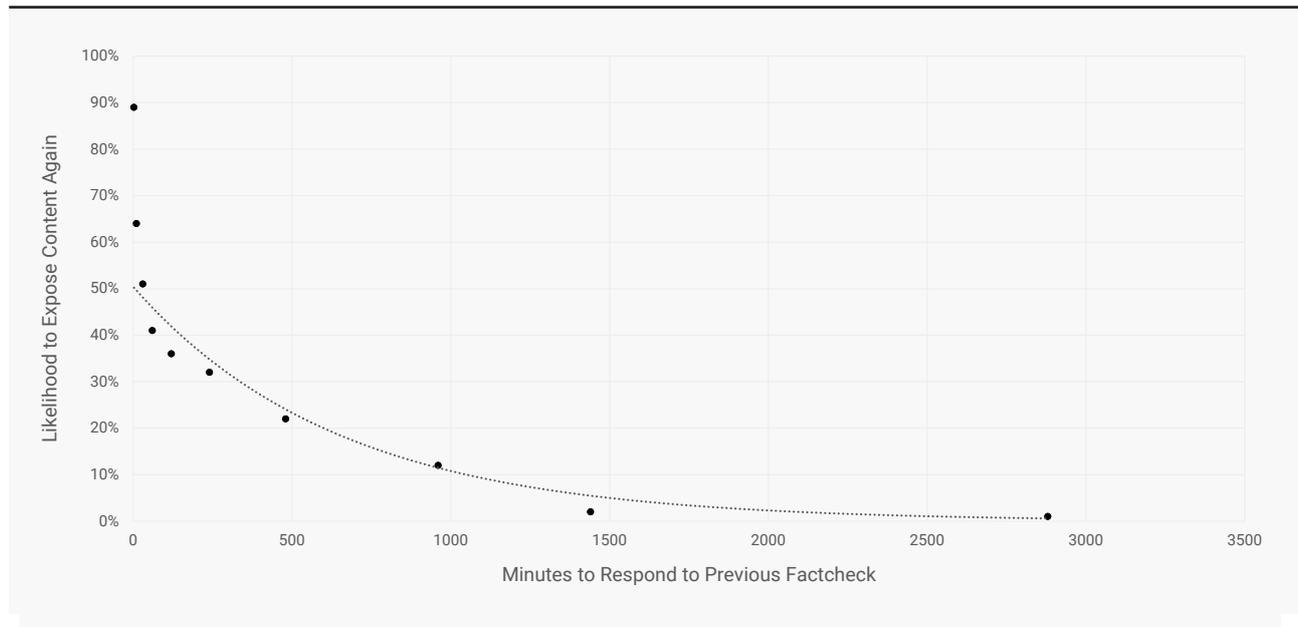
User Requests Answered - Including Key Dates



By looking at the key dates, it's clear that these are days of heightened traffic with 1823 requests answered on 23rd April

Users are more likely to expose content if they are confident they will receive a response - an outright debunk, verification or even a statement acknowledging the nuances and complexity of a particular message.

Impact of Response Time on Verification Frequency



There are persona-wide differences in the sensitivity to response time. However, the general trend holds. If a fact-check is responded to any later than 30 mins over 50% of users are unlikely to expose questionable content.

The immediacy of response appears to help in exposing novel content. Only 793 of all received requests were related to third party fact-checks indicating this methodology uncovers new information previously hidden from fact-checkers and other activists. This phenomenon wasn't just to niche content but also some of the most popular claims. Only 9 of the 25 most popular requests we received were also fact-checked by a third party.

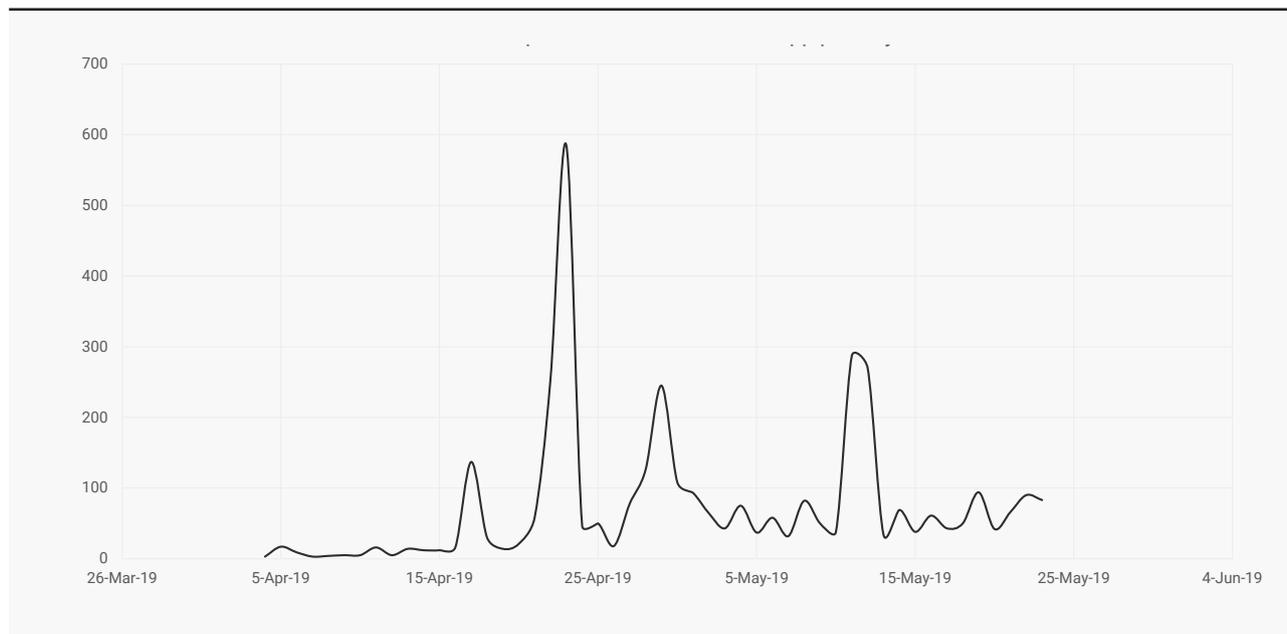
We will be continuing with this body of work focusing on more precisely quantifying the critical mass of active forwarders required in our network to expose a statistically significant percentage of all malicious and false content spread on WhatsApp.

Disseminating Verifications to Audiences Exposed to Mis/Disinformation

The response to each request we received was in the form of a sharable image, similar to the concept used by Verificado in Mexico.

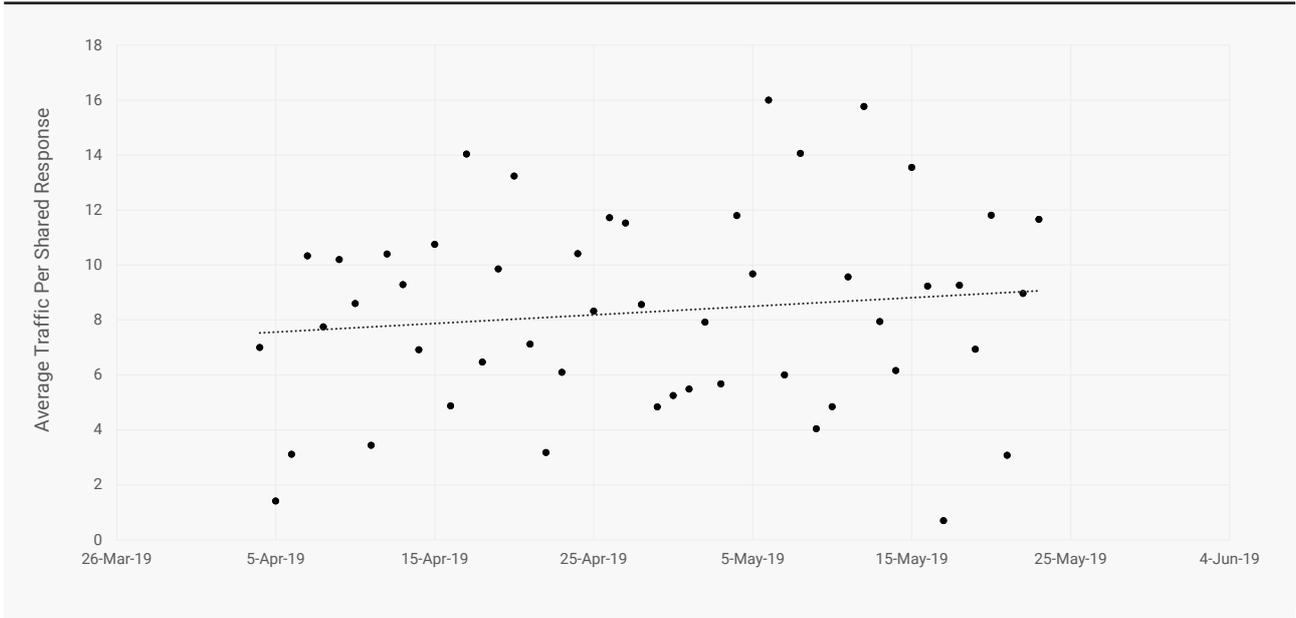
Out of the 11560 User Requests Answered, 3690 were shared to WhatsApp users. The sharing rate was particularly high during key election dates.

No. of Responses Shared on WhatsApp per Day



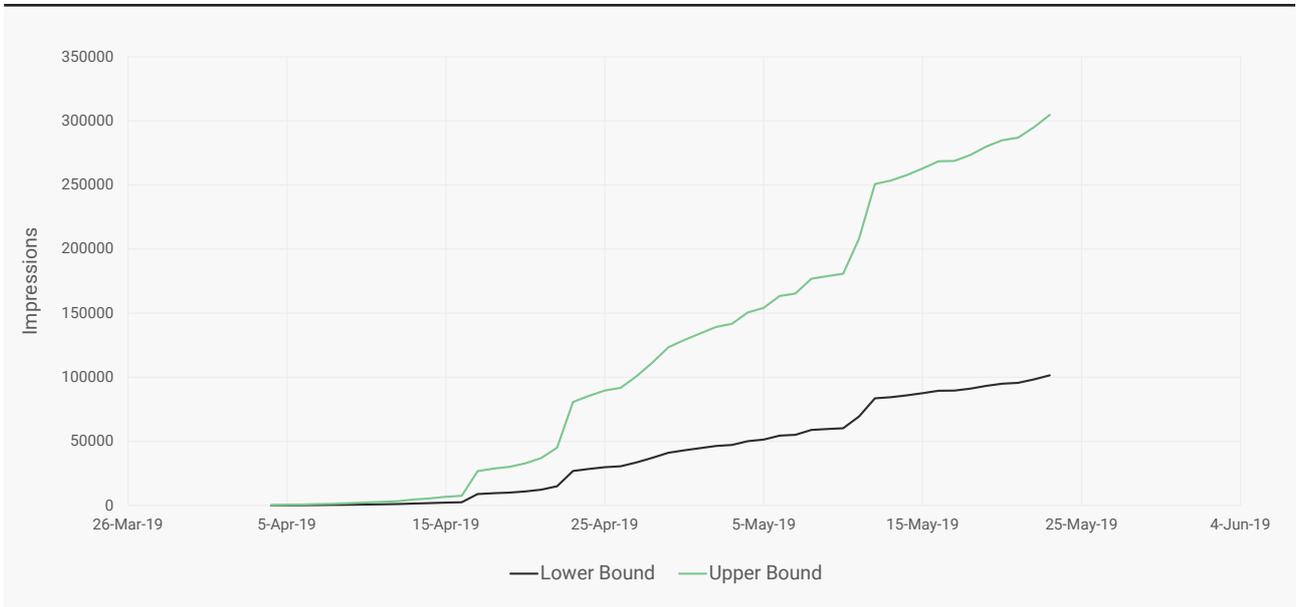
Each response was accompanied by a unique trackable link and unique visits to this URL were logged. The chart below shows the number of unique visits per shared response.

Unique Clicks per Shared Response Factchecks and Verifications



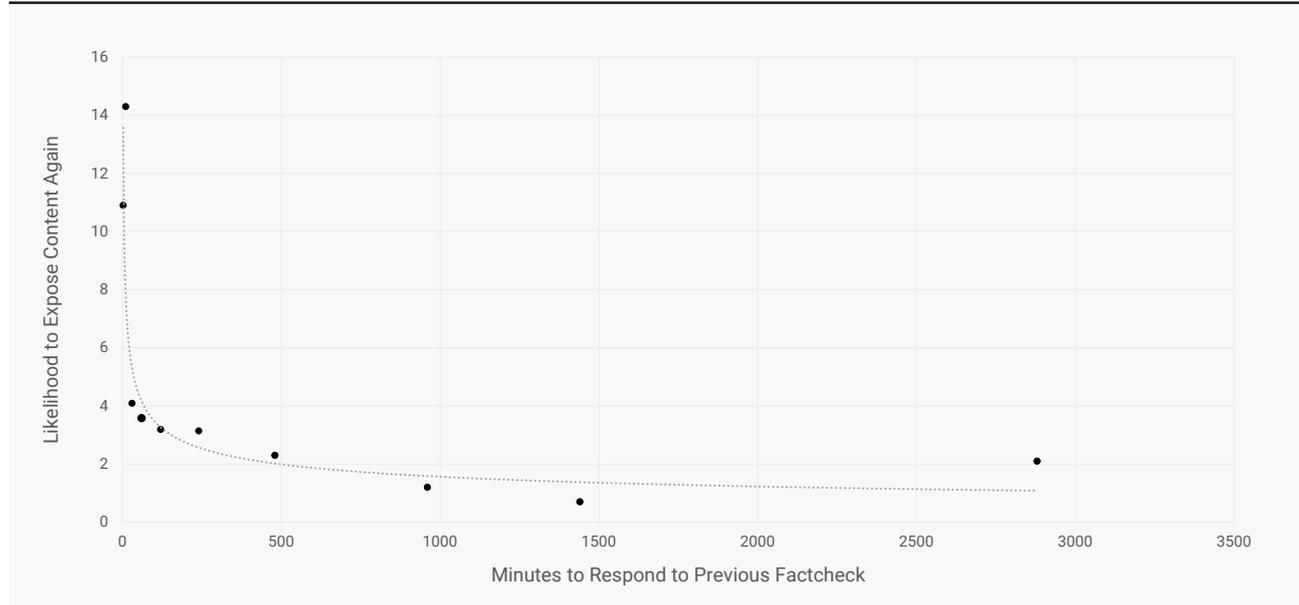
On average the link accompanying the responses received 8 visits. However, the sharable image is designed to be self-sufficient and most viewers are unlikely to click on the link. The clickthrough rate on such messages is expected to be between 10-30% [Source: Gupshup] Using this assumption, we can project the following reach for our responses.

Est. Cummulative Impressions on WhatsApp



The share worthiness of response appears to be highly sensitive to the response time. Automated fact-checks [less than two minutes] receive 11 clicks during their lifetime but are outperformed by fact human fact-checks - these receive 14 clicks during their lifetime.

Impact of Response Time on Expected Lifetime Impressions



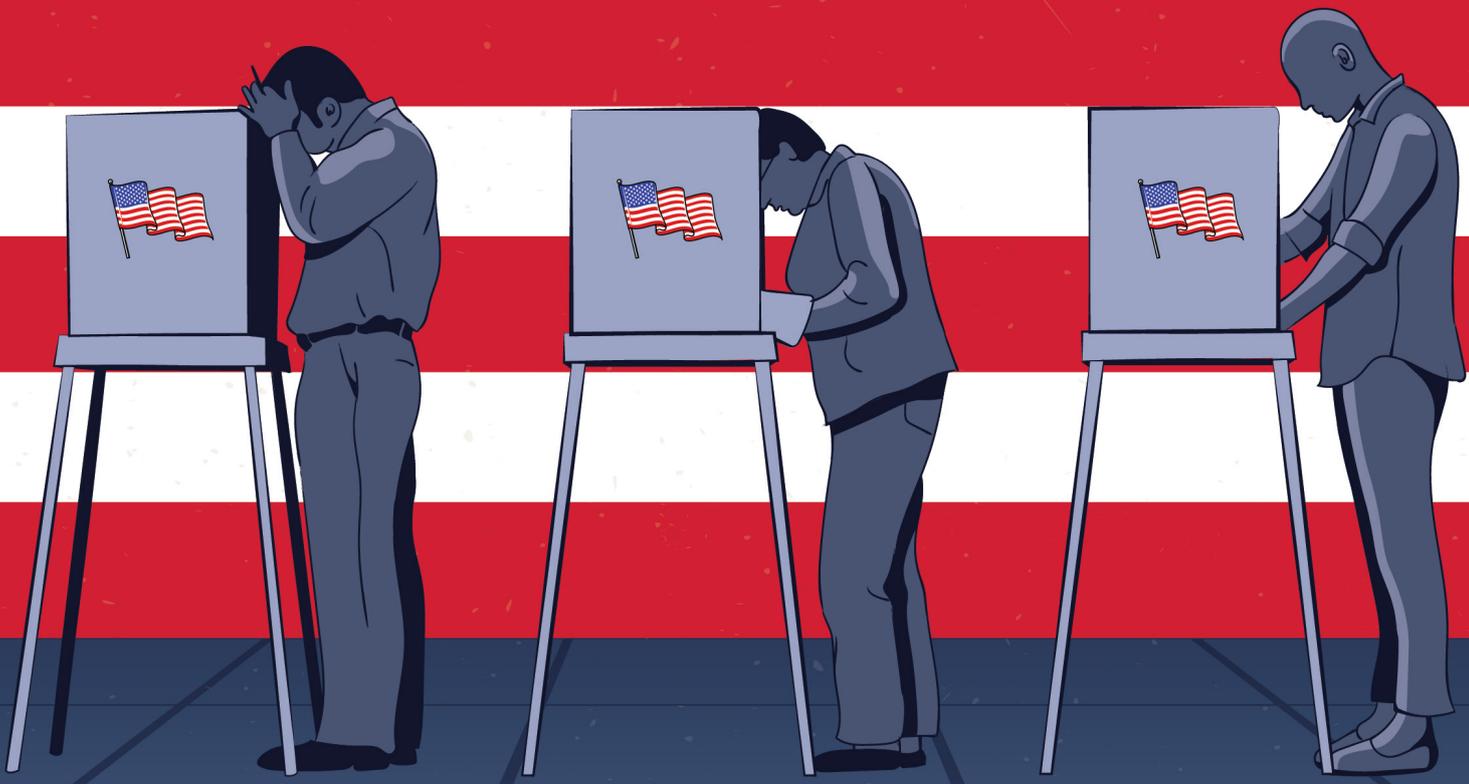
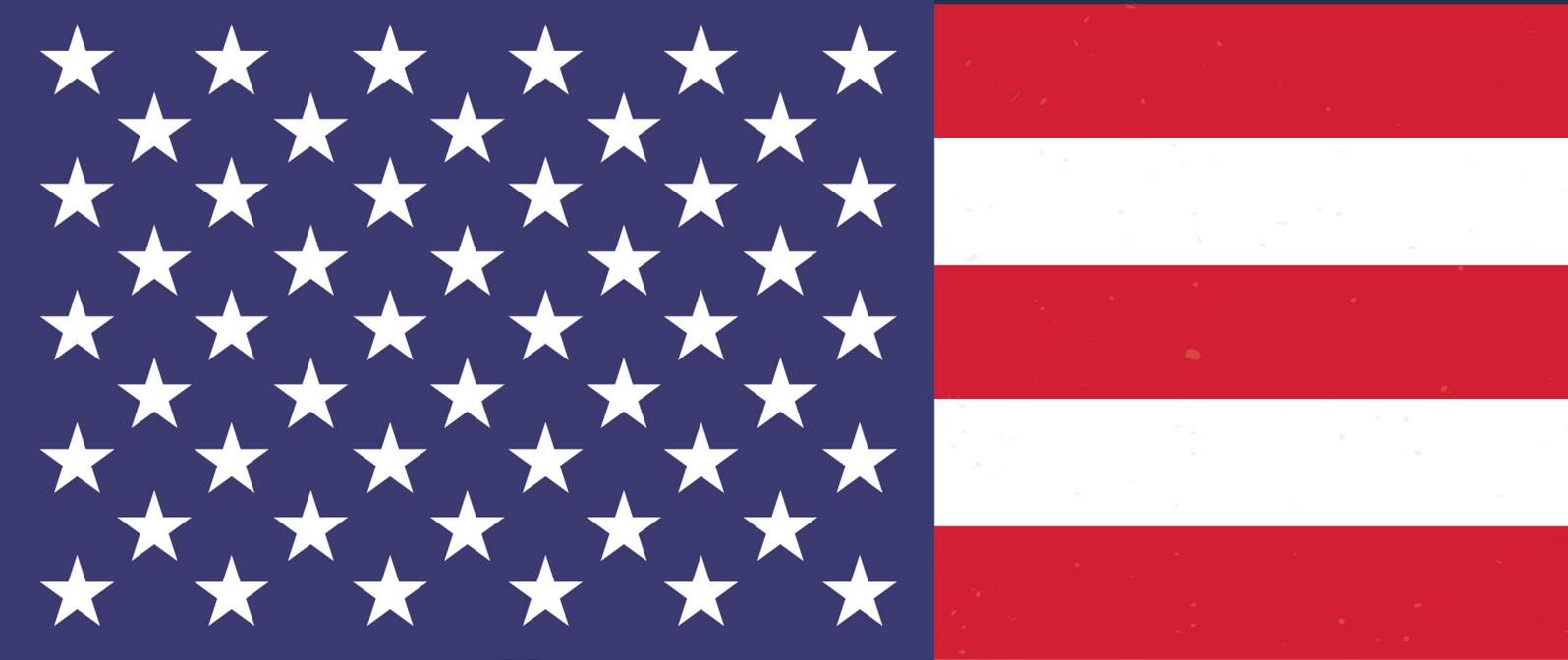
The additional benefit of quick response times would be to slow down or altogether prevent forwarding chain reactions. Based on user surveys, users are almost certain to share verifications of a claim if they discovered it in a group. Based on the link tracking data, our earlier hypothesis that a verification can be forwarded repeatedly has been validated. Furthermore, in a quick turnaround scenario, fact-checks would be delivered to multiple groups that have only recently received the initial problematic message - thereby acting as a deterrent to anyone considering forwarding the original message. There is some evidence to suggest that individual within these groups share the fact-checks more broadly, however, the precise individual and group dynamics resulting from a user sharing a fact-check in their conversation remain unclear.

The average active user of Logically's verification services was able to spread verifications of content to 10-30 WhatsApp users. By scaling these shallow networks nationally to 1-3 million engaged whistleblower/shares, this approach may be able to reach all users affected by mis/disinformation on WhatsApp in India. This approach need not be limited to WhatsApp and in principle can apply to other closed networks such as Facebook Groups, Messenger and Telegram.

An additional hypothesis that can be proposed on the basis of some of these findings is that repeated exposure to quick verifications of content from a single source would build resilience within that group and the routes travelled by both the original content and its corresponding verification. In addition to this, we will be continuing with this body of work focusing on gamification strategies can be leveraged to help incentivise greater sharing rates, and quantifying the critical mass of active forwarders required in our network to disseminate verifications and debunks to a statistically significant portion of audiences initially exposed to the original messages.

08

**Future in
Elections**



Building Video Capabilities

The recent transformation in the digital news landscape has seen online video content become a popular means of accessing news, particularly among younger, more technologically-savvy consumers.

This trend is particularly pronounced in the USA. Reuters' 2019 Digital News Report showed that 65% of US consumers access video news content at least once a week, a figure 7% higher than the previous year and 19% higher than that derived from a UK audience²⁷. Clearly, short-form video news is on the rise and a particularly important feature of the US news landscape.

Video Content and Misinformation

Whilst there is nothing inherently troubling about the migration of digital news consumption from text-based to video format, it is notable that the same Reuters report looks into where exactly this news is being accessed. 52% of respondents access video content offsite - meaning through a third-party aggregator or, more typically, a social media platform - where contextualising information such as the publisher or its sources are far less transparent.²⁸

Deepfake Video Content

A recent public service announcement by American comedian Jordan Peele, posing as Barack Obama to deliver a warning about deepfake technology and its potential impact, highlighted the considerable advancements made in this technology. The video begins as a normal public address, before revealing the comedian as the real source of the content.

Whilst satirical, this serves as a harrowing demonstration of the ease with which a malicious actor could harness the technology to exert ideological or political influence. A 2017 study conducted by researchers at the University of Washington succeeded in a similar deepfake, training machine learning algorithms on hours of footage from Obama's weekly addresses to create an extremely convincing, but completely fake, public address.²⁹

The good news is that the media focus on deepfake technology – and the intense public interest it garners – has meant that technologies being built to detect manipulated video have simultaneously improved. It is highly likely that detecting deepfakes will not, in fact, be the future crisis that many have recently been suggesting. Like with text-based misinformation, the problem may not be so much the verifying of the content itself, but communicating debunks to wide enough audience.

27. 'Digital News Report 2019', Reuters Institute for the Study of Journalism, 2019.

28. IBID

29. Supasorn Suwajanakorn, Steven M. Seitz, Ira Kemelmacher-Shlizerman; 'Synthesizing Obama: Learning Lip Sync from Audio'; University of Washington, 2017.



The Logically Approach

We've embarked upon an approach wherein we would simultaneously assess the credibility of video content and the veracity of specific claims within it. Initial development will focus on building a framework capable of structuring video data; just as our claim detection technology structures unstructured text content for analysis, we need to deconstruct video content to enable automated analysis of isolated elements within it. We've also developed a three-pronged approach which conducts the following analyses in ascertaining the broader credibility and intent behind video content:

- **Authentication:** Evaluation of individual video components to detect alteration or inconsistency.
- **Deception:** Detecting deceptive elements within the video, such as clickbait, misrepresentation or dissonance.
- **Provenance:** Analysis of the video's source for any indications of unreliability.

Domain and Skill Specific Segmentation

As Logically grows and its interests diversify - with new markets and products being pursued - we'll encounter several challenges in scaling our fact-checking operations. New markets will impose different requirements upon our team, with local political, social and media knowledge essential to high-quality fact-checking. Similarly, as we pursue B2B business verticals, our veracity assessment capabilities must be underpinned by relevant industry expertise. This will be particularly relevant for verticals such as Health, where fact-checks must be carried out by qualified GPs.

The challenge in scaling our fact-checking team will be in maintaining the significant advantage we've gained in terms of operational efficiency whilst integrating regional and domain expertise into our processes. Our strategy for scaling the team to meet these requirements treats the two issues - local and domain expertise - as separate requirements, with a specific approach taken to each:

Regional knowledge for new B2C markets

In addition to building out our Indian fact-checking team to support a greater volume of fact-checks, Logically will establish regional satellite teams responsible for the moderation of our Indian team's output, ensuring that they accurately reflect the local socio-political climate and conform to any regional fact-checking standards or consumer expectations before publication. This will enable us to maintain a cost-effective B2C fact-checking operation, with a single, well-trained team responsible for all fact-checking, and more extensive local moderation procedures to safeguard the quality of our fact-checks.

UK and US Elections

While analysis into the events that surrounded the 2016 US presidential elections will continue for years to come, what is clear is that misinformation that circulated in the lead up influenced the outcome of the election itself. Having observed this phenomenon, as well as the accompanying hyperpartisanship and break down in civil discourse that accompanied it, Logically set about designing its news app to try and address these externalities within the information ecosystem.

In particular, we designed a fact-checking operation specifically for users, so that they could request verification and clarification of things that they were unsure or suspicious of, receiving an easy-to-understand response that could be shared quickly and easily with their friends and families. We believe that this user-centric model of verification can help to address some of the information disorders that became apparent during the 2016 election, by providing people with a simple and quick way of accessing objective, apolitical facts about the statements and claims that they've heard, and curbing the flow of questionable information organically.

Fast forward to the 2019 Lok Sabha elections in India, the largest democratic event ever to take place, and Logically had evolved from just a late-night thought to a full-blown startup with several products in the process of development and launch. Given the prevalence of misinformation in India, especially over closed messaging platforms, we felt that we were ready to tackle our first election.

Our fact-checking team learnt some invaluable lessons during their work on the Indian elections. We've since tailored our verification methodology to prioritise high-quality, primary sources more than ever before, establishing libraries and databases to facilitate our work. We also noticed the importance of speed, given the fast-paced nature of the political news cycle in the run up to an election. We've since expanded our team by a further five individuals, and are working on building an extensive claim library to support the capacity of our automated fact-checking efforts.

Looking forward to next year's US presidential election, and the possibility of a general election in the UK in the near future, we hope that these lessons can be applied by our team to address misinformation around elections throughout the world. Our team is working hard to familiarise themselves with the local political contexts in these locations, developing resource libraries and source taxonomies to facilitate rapid, high quality verification. We are also preparing and training to fact-check live events so that we can address false or misleading claims made by politicians at their source, and provide the population with better information with which to inform their decisions.

09

Appendix

Logically. 264511af
24/09/2019 10:09 IST



FALSE

Viral picture of dead cows is from Rajasthan, where cows die due to starvation

The report published by Daily Nation and Eating My Ethics article of Australia refutes the statement.





Logically. eba8728d
19/04/2019 23:04 IST



FALSE

Sonia Gandhi is richer than the Queen.

Contradictory information from the BBC and debunks by factchecking organisation disprove this claim.





Logically. 49a6983e
19/04/2019 17:04 IST



FALSE

Deepika Padukone was born in Copenhagen, Denmark and doesn't have an Indian passport.

A photo of the Passport on Deepika's verified official Instagram account and her interview video is inconsistent with the claim.




Logically. 333aee28
22/05/2019 12:05



PARTLY TRUE

Narendra Modi visited West Bengal 17 times during the election campaigns for 2019, in an effort to make inroads with his Hindu nationalist agenda, provoking sporadic violence and prompting the Election Commission to cut off campaigning there.

An ECI order confirms the campaign ban in West Bengal but the cause of the violence was Modi's agenda or not is unverifiable as it is an opinion.





Logically. 8a0ee73d
21/06/2019 14:05



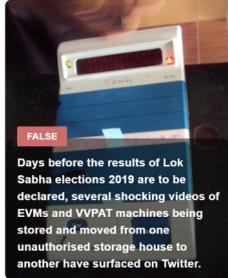
FALSE

The Pulwama attack of 14 February 2019 was a planned conspiracy of Bharatiya Janata Party and a fake attack on Pakistan.

The official press releases by the Indian Ministry of External Affairs is inconsistent with the statement.




Logically. 545a30a1
22/05/2019 14:05



FALSE

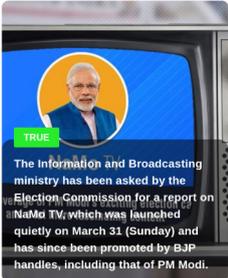
Days before the results of Lok Sabha elections 2019 are to be declared, several shocking videos of EVMs and VVPAT machines being stored and moved from one unauthorised storage house to another have surfaced on Twitter.

An official press release by the Election Commission of India (ECI) is inconsistent with the statement.





Logically. 582a12c9
11/04/2019 12:04 IST



TRUE

The Information and Broadcasting ministry has been asked by the Election Commission for a report on an aMa TV, which was launched quietly on March 31 (Sunday) and has since been promoted by BJP handles, including that of PM Modi.

The claim is proven to be True, as it is consistent with the tweet from Narendra Modi and BJP's official verified Twitter accounts.







LYRIC JAIN | CEO, LOGICALLY

 lyricjain

www.logically.co.uk | contact@logically.co.uk

 LogicallyHQ

 LogicallyHQ

 logically.hq