

Catch, Corroborate and Communicate: The Contours of Professional fact-Checking In India

ARSHIA ARYA, Microsoft Research, India, India

SOHAM DE, Microsoft Research, India, India

SALONI DASH, Microsoft Research, India, India

JOYOJEET PAL, School of Information, University of Michigan, India

The work of professional fact-checking involves three broad processes - selection of the issues that need to be fact-checked, verifying a story that is flagged as being a falsehood, and finally disseminating a fact-checked version of a story, each of which have unique challenges. We conducted in-depth interviews on their professional practice with fact-checkers of five major organizations that work in debunking stories daily, alongside conducted a large-scale analysis of the archives of fact-checked articles and studied the social media footprint of fact-checkers from these organizations. Our findings show that the outcomes of work with each of the three processes influence what stories get debunked by organizations, how much public impact they have, and the depth of debunking in each story, all of which in turn play a role in the scope and breadth of individual organizations' impact. Our findings have implications from both the supply side of how organizations can place themselves within the fact-checking ecosystem, and from the demand side of how consumers can consider the work and positioning of their preferred fact-checkers.

ACM Reference Format:

Arshia Arya, Soham De, Saloni Dash, and Joyojeet Pal. 2023. Catch, Corroborate and Communicate: The Contours of Professional fact-Checking In India. In . ACM, New York, NY, USA, 24 pages. <https://doi.org/XXXXXXX.XXXXXXX>

1 INTRODUCTION

Fact-checkers have risen to the fore as a key part of the global information environment as the democratization of information creation and viral dissemination has given rise to a massive increase in unverified and often explicitly malicious information. They are also by the necessary scope of their work the set of professionals most likely to offend someone with every single story they do, since their work is essentially combative in nature, taking down stories that someone has chosen to push.

Professional fact-checking is a multi-step process. First stories need to be selected for fact-checking. This involves stories being reported to a factchecker. There are a range of specialized tools that allow for this, in addition to the flagging of stories through social media. This process can thus be intense and overwhelming, which leads to the related challenge of selection – prioritizing. A fact-checking organization must decide which stories it would like to focus on, which it can competently debunk if needed, and which are not within its priority domains.

The second step in professional fact-checking is around verification. Once a story has been selected for fact-checking, it needs to be checked through various technical and non-technical means. One part of this is the format of the story, and whether this allows for easy digital verification – such as if there are images, audio or videos that can be reverse-searched, purported facts that can be confirmed through an online search process through string matches. An important element

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2023 Association for Computing Machinery.

Manuscript submitted to ACM

53 of the verification process is the expertise involved in the verification process. In some cases, the core fact-checking
54 team has specific skills or past expertise that make it easier to factcheck a certain story, especially if it has particular
55 technical content that require careful examination. Related to this is the issue of context, fact-checking a story may
56 need access to a region or specific type of professional, which may not be easily available.
57

58 The last step is that of publication. fact-checkers are mindful of the audiences they write for, and frame their stories
59 accordingly. This can have implications around political preferences of the audiences or the stories they rely on a specific
60 factchecker for. The second key aspect related to publication is that of dissemination, in which the factchecker needs to
61 think about what are the specific strengths of their organization in terms of online reach. While certain fact-checkers
62 are part of large organizations with significant mainstream media following, others rely on the reach of individual
63 fact-checkers or their handles' social media footprint. This matters for what kinds of stories get covered, and what
64 stay in the minds of the audiences they reach. This also means that when a fact-checking organization takes a certain
65 position against a particular kind of misinformation, say right-wing misinformation, or anti-science misinformation, its
66 dedicated audience expects more of the same, influencing what gets debunked and how.
67

68 fact-checking as a profession is here to stay. While its contours may change over time, the steady spread of
69 misinformation suggests that there will be a space for professionals to systematically examine the veracity of stories
70 that have a potential for wide impact. The fact-checking ecosystem exists within several these include the decline of
71 trust in the news in India, with corresponding increase in access to the news through non-mainstream sources [9]. This
72 follows similar patterns worldwide, where professional journalists report a drop in public trust in the mainstream news
73 worldwide[82].
74

75 While fact-checking is growing, it is both expensive, and risky work. This has meant that the audience is a major
76 driver of who gets into this work. Professional fact-checking tends to be a largely urban activity, and driven by the
77 stories and interests of its audiences. This impacts the pipeline of not only what gets reported to fact-checkers, but
78 furthermore what they pick to debunk, what their staffers are best equipped to debunk, and what, of their debunked
79 stories, once put out in the general public, gets the most affective appeal. In a nutshell, this covers the spread of selection,
80 verification, and dissemination in the professional fact-checking ecosystem.
81

82 In this work, we frame fact-checking within these constraints of how fact-checking is organized. We aim to address
83 three questions in this work.
84

85 **(RQ1)** What are the challenges in the life cycle of fact-checked articles in the Indian context?

86 **(RQ2)** What are the challenges faced by the fact-checkers in the process?

87 **(RQ3)** How do these challenges affect the coverage, reportage and dissemination of these fact-checks?
88

89 The fact-checking process and the impact it has on the society has been of great interest to the CSCW community.
90 Recent papers have done in depth study on fact-checking infrastructures [45], the work practices of fact-checkers [59]
91 and the overall role they play in combating misinformation in the context of the Global South [36].
92

93 Our work focuses on the challenges in the fact-checking pipeline on each step, in the Indian context. India has one
94 of the largest number of IFCN certified fact-checkers, with many independent fact-checkers helming the movement
95 of fact-checking. While falls weak [47] in the face of the viral spread achieved by closed-group platforms [76] and
96 forwarding enabled messaging platforms Coupled with this, Indian misinformation landscape is laced with affective
97 elements, aiming at the political, religious, financial sentiments of people, which makes fact-checking a far more
98 nuanced.
99

100 An important part of the ecosystem is the bad actors who explicitly engage in attempts to muddy the information
101 environment, alongside their followers, who may or may not have explicit malintent, but trust the channels through
102

105 which this information flows. This is particularly true for India, where a polarized political environment has impacted
106 the supply- and demand-sides of misinformation.

107 Our contribution in this work thus are as follows:

- 108 • We present a systematic study of the fact-checking pipeline in professional organizations in India, arguably the
109 largest open market for broadcast media consumption.
- 110 • We present a mixed methods study, using both in-depth interviews with fact-checkers and evidence from large
111 scale data analysis from the fact-checking archives to map which organizations engage in how much activity,
112 on what domains.
- 113 • We examine the second order effects of social media outreach on fact-checking by studying the role of social
114 media and mainstream media in the work of fact-checkers.

115 In summary, the key contribution of this work is a rich description of both what the fact-checkers are dealing with in
116 their daily practice, and the caveats with which we need to understand their work. The specificities of the Indian context
117 in which this work takes place also underlines the importance of the quality of the media environment and polarization
118 in society, all elements that frame both the believability of information and the acceptance for fact-checking.
119

120 2 RELATED WORK

121 2.1 Social Media and News Consumption

122 Social media platforms revolutionized the pre-existing models of information and news consumption [25]. Firstly,
123 by taking over the role of distribution of news [61] and secondly, giving active control of this distribution to large
124 technology companies. And thirdly, providing a model where the users of these platforms have a high-choice media
125 environment with free participatory access to various forms of media [62]. This has given rise to the double-edged sword
126 of media pluralism [77]. On one end, the increasingly digital media environment has provided the users opportunities
127 of diverse sources and perspectives, but on the other end, the dominance of very few large players in the environment
128 pose a concern for diversity in news production. The concentration of power allows influence over the content of public
129 discourses, with their owners' private interests frequently determining the topics they cover .

130 From Arab springs [83] to the suicide of a popular film star[2], the world has witnessed the influence of power con-
131 centrations. Additionally, these temporal events bought to light the nature of technological infrastructures in promoting
132 and impeding invested interests [10]. While literature on pluralism argues that the technological infrastructures are
133 neutral [75], the systems that allow functioning of these platforms are owned by companies with personal agendas and
134 ideological preferences.

135 Problems that social media platforms pose for news consumption [26] and dissemination. Several limitations of social
136 media like the presence of filter bubbles [71], having algorithm driven engagement [50] through influencer networks
137 and the need for content moderation [29] make it a breeding ground for contrived theories and amplification of blatant
138 lies.

139 For citizens, the move towards an increasingly digital, mobile, and social media environment represent the de-
140 velopment of a more high-choice environment in most respects—though there is less diversity in terms of original,
141 professionally produced news on some issues and areas, especially locally. [40]. With increasing concerns on pluralism
142 and the declining credibility and trust in the media sources [20], there are several challenges with the trust in media. As
143 a consequence, this leads to problems of media manipulation, disinformation [53], and propaganda. The consequences
144

157 for informational diversity and media pluralism [15], key to freedom of expression [43], understood as the ability to
158 impart and receive information are mixed.

159 The presence of news consumers on social media is ubiquitous [74]. The key feature that makes social media stand
160 out is the direct communication to users [65]. The baton of influence on content of public discussions has been passed
161 to the 'digital influencers' [13]. by attracting politicians, media houses, businesses, governmental institutions and
162 other knowledge institutions onto the platform [6]. All this allowed removal of one layer checks and balances [3]. The
163 development is continuing with the rise of "distributed content", where news is not only found via, but consumed on,
164 platforms controlled by third-parties that do not produce their own content[46].
165

166 With shift of political communication and news consumption onto SM platforms, the affordances of social media
167 platforms [12] (features, policy etc) has changed quite a bit. Verification of information has become a problem, crippled
168 epistemology [54], citizen journalism [4], declining trust in institutional actor, echo chambers increasing polarization
169 [69], all of these aspects are creating a vortex, which constantly is asking for verification of information. With these
170 convolutions in place, the question of authenticity becomes pressing, leading to the emergence of fact-checkers for
171 addressing the problems with verification of information. But who are the verifiers? the 'elites' or the influencers
172 because policy of social media platforms fall short in contextuality and disproportionately towards a favoured few. Our
173 findings are also suggesting same: that even though IFCN certified fact-check organizations exist, the influencers of
174 those organizations receive more engagement. This consistent ambiguity has put fact-checking organizations on a
175 pedestal. Our work critically analyses their process of verification.
176

177 There is policy intervention required for media an disinformation literacy [42] did not move as quickly as increasingly
178 digital, mobile, and social media environment which perpetuates fake news [52]. The biggest platforms are increasingly
179 important for the distribution of news [67], but invest little or nothing in news production.
180

181 2.2 Rise and Role of fact-Checking

182 The role of fact-checking as an institution and fact-checkers as individuals have seen multiple revisions in the past
183 century. The turn of the 20th century paved the way for fact-based media among the prevalent culture of yellow
184 journalism in the 1980s [23]. The subsequent professionalisation of the media led to a concerted attempt to codify ethics
185 and a greater focus on ideals of accuracy and unbiased reporting - the setting up of the Bureacracy of Accuracy and
186 Fairplay by Ralph Pulitzer being an illustrative example [21]. fact-checking in journalism emerged as a pre-publication
187 process to scrutinise facts in articles before they were published. This is around the time when the term *fact-checking*
188 was introduced in an advertisement for a newly set up research department at the Time magazine. From an objective
189 scrutiny of every word to be published, fact-checking gradually emerged as a larger exercise to ensure an aggregate
190 truth emerging out of objective facts. This broader responsibility on the fact-checking process led to a blurring of roles
191 between the factchecker and the journalist, marked by a noted decline in the fact-checking jobs.
192

193 The resurgence of fact-checking was driven primarily by American politics and led to the rise of independent, external
194 fact-checking organisations, as opposed to the ante-hoc verification of facts in the past [34]. US Presidential campaigns
195 focused on misinformation such as that of Reagan's is often considered to be a major impetus of the boom of post-hoc
196 factcheeking in American journalism [22][36] as were checks on political advertisements (*Adwatch*) [28]. The early 21st
197 century saw the emergence of Pulitzer-winning fact-checking organisations such as *PolitiFact* an *FactCheck.org*, which
198 played a major role in popularising fact-checking well beyond American shores. This evolution of fact-checking from
199 research departments of American media houses to external stand-alone institutions has been well studied in a series of
200

209 prior works [31, 33, 35]. At the time of writing, Duke Reporters' Lab [51] records the global number of fact-checking
210 organisations at 391 - all of which have surfaced within a brief span of 18 years.

211 Investigations into fact-checking processes, assistive and interventional infrastructures and the invisibility of human
212 infrastructures has been of particular interest to the CSCW community. In particular, prior work [45] has outlined the
213 role of human and technological infrastructures within organisations as well as its adaptation and perception in the
214 global south [36]. While academic interest around fact-checking remains to be centered around the global west, there
215 are numerous studies [24, 60, 81] that highlight various nuances of fact-checking in resource-constraint environments.
216 On the issue of publication and dissemination, recent CSCW studies have addressed most on-ground dissemination via
217 informal networks and public dialogue [48, 73] as well as spread in online platform-driven spaces [27, 41]. Only few
218 studies [78] have presented analyses of historical fact-checked articles as a window into the outcomes of fact-checking
219 processes and we aren't aware of any such work positioned in the Global South.
220
221

222 We extend [45]'s exhaustive study of factchecking infrastructures by
223
224
225

226 3 BACKGROUND

227 3.1 Misinformation Landscape in India

228
229 3.1.1 *ICT as an enabler of misinformation.* Misinformation as a global challenge has been the focus of many recent
230 studies within the CSCW (Computer-Supported Cooperative Work) community [5, 7, 37, 44]. One may attribute a rising
231 interest in misinformation studies to recent events in the global west including notably, the Russian interference in
232 2016 US Presidential elections [7] and the Russian Internet Research Agency (IRA)-led efforts to undermine social BLM
233 protests by cultivated radicalised communities on Twitter to interfere with the movement [72]. Across the Atlantic,
234 studies [37] have often focused on targeted bot activity around Brexit conversations online.
235
236

237 The Global South has a democracy distinct from the Global North. Consequently the nature of misinformation in these
238 communities often take advantage of religious sentiments and limited truth-verification infrastructures of low-income
239 communities. Many recent studies [8, 19, 38, 63, 68] within the ICTD and CSCW communities have underlined the said
240 religious undertone in misinformation prevalent in the Global South. We reemphasise that such misinformation in the
241 Global South can often lead to serious consequences including lynchings [79] and infodemic-caused deaths [39].
242
243
244

245 3.2 Mitigation of Misinformation

246
247 Efforts to mitigate misinformation have gained importance at scale from initiatives to stop misinformation right at the
248 source using content and platform moderation [30] practices. There are several initiatives by platforms to implement
249 policies that exacerbate content moderation and flag problematic content. For instance, Twitter takes a proactive stance
250 on purging suspicious accounts with a high probability of exhibiting social bot-like behaviour. Media literacy is a
251 proactive and preventative approach to deal with the constant influx of claims that float on the digital media. However,
252 this is a poor strategy for low literate users [1]. Citizen journalism [80] is another approach that is a key pillar of
253 practical addressal of misinformation, especially in close knit communities and low resource settings. Other stakeholders
254 like Investigative journalists also take charge of debunking false claims online [55]. Among these approaches, fact-
255 checking emerges as a regulated, organised approach to systematically take an active approach towards mitigating
256 misinformation.
257
258
259
260

Table 1. List of Interviewed fact-Checkers

P#	Role within Organisation	Gender	IFCN Certification
P1	Lead, Factcheck Team	Male	Yes
P2	Founder and Journalist	Male	Yes
P3	Founder	Male	Yes
P4	Subject Expert	Female	Yes
P5	Senior Editor	Female	Yes
P6	Co-Founder	Male	Yes
P7	Factchecker	Female	Yes
P8	Co-Founder	Male	Yes
P9	Managing Editor	Male	Yes

4 METHODOLOGY

We use both qualitative and quantitative research elements in the study to map and unpack the process of fact-checking and the environment it inhabits. We use expert interviews of fact-checkers in India as our primary source of data. The qualitative data collected provides us insights into methodology of fact-checking process, relationship with social media platforms, impact and consequences of misinformation and probable solutions to combating misinformation. We then use quantitative data, to corroborate of primary findings through data visualization. The quantitative data used in the study comes from an archive of fact-checked articles and the twitter interactions of fact-checking organizations in India. The exploratory nature of the study provides an in-depth understanding of real-life contexts and issues and how they unfold in practice, that can better inform the future of work in the field of social computing. The sections below elaborate on our data sources, including the sampling of data, questionnaire design and analysis.

4.1 Expert Interviews

We conduct expert interviews (N=9) with leading fact-checkers in India. We sampled a list of active fact-checking organizations in India from the Duke Reporter's Lab [51]. As there is no single definition of fact-checking organizations summarizing their basic features, we adopt the following definition by International fact-Checking Networks's (IFCN), 'organizations that regularly publish nonpartisan reports on the accuracy of statements by public figures, major institutions, and other widely circulated claims of interest to society', for the purposes of this study [66]. The association also has a code of principles to certify fact-checkers [70]. At the time of the data collection there were 11 IFCN certified fact-checkers in India. We contacted the individual heads of every major fact-checking organisation in India via their official email and used convenience sampling [84] to curate the final list of participants for the interview. The email contained a brief description of the project and the participant was given an option to take the interview either in-person or over any video-calling platform of their choice. Our last author has established connections and networks within major fact-checking organisations in India and we leveraged their reach to solicit as many participants as possible. If the personal email of any individual wasn't available, we tried contacting the organisation seeking interview with the said individuals. Our assumption in contacting heads of organisation was an inherent hierarchy in the organisation and to seek information about the approach of the organisation as a whole, as opposed to the perspective of individual fact-checkers in composing factchecks. The study received ethical approval from xx, India (hidden for blind-review).

The questionnaire for the interviews was designed with an exploratory study in mind. The nature of the questions are open-ended, to gather details and encourage elaborate and deeper evaluation of fact-checking as a process. Based

on recommendations of studies related to social media communication and misinformation [11, 32, 49, 58?], we include questions under the broad theme of methodology of fact-checking process, relationship with social media platforms, impact and consequences of misinformation and probable solutions to combating misinformation. Our questionnaire provides us insights into the nature and dynamics of false information, practices of fact-checking, everyday interactions with technological aid, perceptions on targeted audiences, views on technological and legal methods of combating misinformation and experiences of factchecker in the volatile sociopolitical environment in India. We piloted the questionnaire with two experts on fact-checking and based on that feedback we revised the questionnaire for language and included questions on fact-checkers' relationship with mainstream media. The questionnaire was then sent for approval to Institutional Review Board (IRB) of the authors' institutions.

The expert interviews were conducted by the first and second authors of the paper either in-person or over zoom. In either case, the interviews required a written consent of participants, in accordance to an approved ethics committee. The interviews lasted from 25 minutes to 90 minutes in length, with an average length of around 45 minutes. Each interview began with a brief discussion about the intent of the project and introduction of the interviewer. This was followed by an enquiry into their motivations and objectives for being in the fact-checking space in India. We follow this up with a sample set of questions and an associated script. Whenever possible, the interviewer nudged the participant to elaborate on details of the fact-checking process, including factors behind prioritisation of stories to check, available technological aid, vulnerabilities of fake news and the overall effect of resource constraints on fact-checking.

These interviews were then open coded by the authors using automated speech-to-text translation software, followed by a manual verification of the translation output. The transcripts were then open-coded deductively by multiple authors independently of each other. The codes were then compared and revised till the authors could arrive at a cohesive set of codes grouped by precise themes to structure the knowledge gained from the interviews, which are presented in subsequent sections as responses to the research questions introduced in this study.

4.2 Archive of fact-checked Articles and Twitter Data

Table 2. **Archive of Published fact-Checks**

Parent Organisation	Domain	Count
News Mobile	newsmobile.in	2266
BOOM	boomlive.in	1082
Factly	factly.in	762
India Today	indiatoday.in/factcheck	590
AltNews	altnews.in	457

4.2.1 Data. While the interviews form our primary data source, we also make generous use of two other data sources - a historical archive of published factchecks (N=6027) by Tattle Civic Technologies under an Open Database License (ODbL), and the tweeting history of the main fact-checking organisations in the study extracted from Twitter's v2 API. The archive of factchecks were pre-processed to retain checks only in the English language. We also focus on published checks and tweets from 5 fact-checking organisations (AltNews, NewsMobile, Factly, BOOMlive and IndiaToday), each of which have published at least 300 factchecks over the period of study (March 1, 2020 to December 31, 2020). In addition to the tweets, we also collect the followers of the organisations on Twitter as a proxy for the social-media audience for their published factchecks.

365 4.2.2 *Analysis*. Prior to analysing the archive of published factchecks, we run certain preliminary filters on the data
 366 to remove all non-english language stories. The final archive has a total of 6027 fact-checked stories from the 5
 367 organisations in our study. Each of these are stories where the claims are based primarily out of India. We perform
 368 basic frequentist and statistical analyses to report general publishing volumes across organisations over the period
 369 of study. We also present wordclouds of the headlines of these articles in our study. These clouds were generated by
 370 calculating frequency of terms contained within headlines. We iteratively remove the most common terms (such as
 371 *fact, check, claim, shared, viral, truth, covid* etc) so as to highlight nuanced differences among reporting styles of these
 372 organisations. We also present certain bag-of-words (BoW)-based analyses of headline and body to distinguish between
 373 media artefact-based debunking from those involving on-ground reportage. Keywords used to detect use of media
 374 artefacts include *old video, old image, photo, reverse image search, yandex* and those for signs of on-ground investigation
 375 include *on-ground, (organisation_name) visited, went to, first hand* and so on
 376
 377
 378
 379

380 5 FINDINGS

381 5.1 News Monitoring and Prioritisation

382 Given the rapid penetration of internet in India and dissemination of News online, the Indian audience is exposed
 383 to a very high volume of information on a daily basis. This is a severe hurdle in the fact-checking process as it in
 384 infeasible to keep up with all information that gets put out there. Thus the process of fact-checking relies on a robust
 385 crowd-assisted monitoring and sourcing pipeline to keep an eye out of suspicious stories that require a check. This is
 386 currently done in two primary ways - (a) algorithmic nudging and (b) crowd-sourcing.
 387
 388
 389

- 390 (a) *Algorithmic Nudging*: Most major fact-checking organisations have access to proprietary tools such as CrowdTan-
 391 gle and Facebook’s algorithmic detection queue (informally referred to as the *Facebook queue* in the interviews)
 392 to flag potentially suspicious news. The drawback of such black-boxed approaches involve a lack a explainability
 393 as well as a potential propagation of biases in algorithmic recommendations. Unfortunately due to NDAs with
 394 Facebook, the participants were unable to divulge exact details about *the queue*.
 395
- 396 (b) *Crowd-sourcing*: Data and interviews reveal that a major chunk of monitoring news is crowd-sourced to the
 397 audience of factchecks. Most organisations have dedicated WhatsApp hotlines to solicit requests for factchecks.
 398 Organisations are also known to frequent certain social media pages on Facebook and Twitter who are more
 399 likely to perpetuate suspicious content onto their audiences. This mode of monitor also allows the organisation
 400 to adhere to their advocacy objectives of preventing maximum harm by proactively seeking to debunk suspicious
 401 information their audience is subject to. In Fig 1, we attempt to study the source of published factchecks across
 402 organisations to uncover potential disparities in reliance over certain platforms over others. Fig 1 shows the
 403 distribution of articles containing mentions of certain social media platforms in the body of the articles. Our
 404 assumption that these platform mentions reflect their monitoring systems relies on the IFCN guidelines that
 405 require an organisation to explain the process of fact-checking in their published stories. We followed this up
 406 by a ground truth analysis to ensure that these mentions were indeed true reflections of the source of these
 407 factchecks. We critically observe that while most organisations rely on Facebook and Twitter for stories to check,
 408 NewsMobile has an over-reliance on WhatsApp for sourcing stories to check. We argue that due to inherent
 409 platform differences, this subjects NewsMobile to an entirely different distribution of stories to monitor and
 410 verify. From Fig 3 we also observe that NewsMobile consistently publishes the the largest volume of factchecks
 411 among all other organisations.
 412
 413
 414
 415
 416

After monitoring comes the issue of *prioritisation*. The decision of an organisation to prioritise articles to check is based on multiple factors ranging from advocacy to resource-limitations. We present the following quote from one of the participants as an illustration:

"One basic criteria that we apply is the impact, if this wrong news keeps circulating what impact will it have on society's general. If we have 5 requests on the table, we select 2 on the basis of which of these will be more relevant and will have more impact." - P1

In addition to potential to cause harm and the impact of suspected falsehood, we argue that resource constraints pose a serious hurdle in the prioritisation decision. We illustrate several facets of this in subsequent sections.

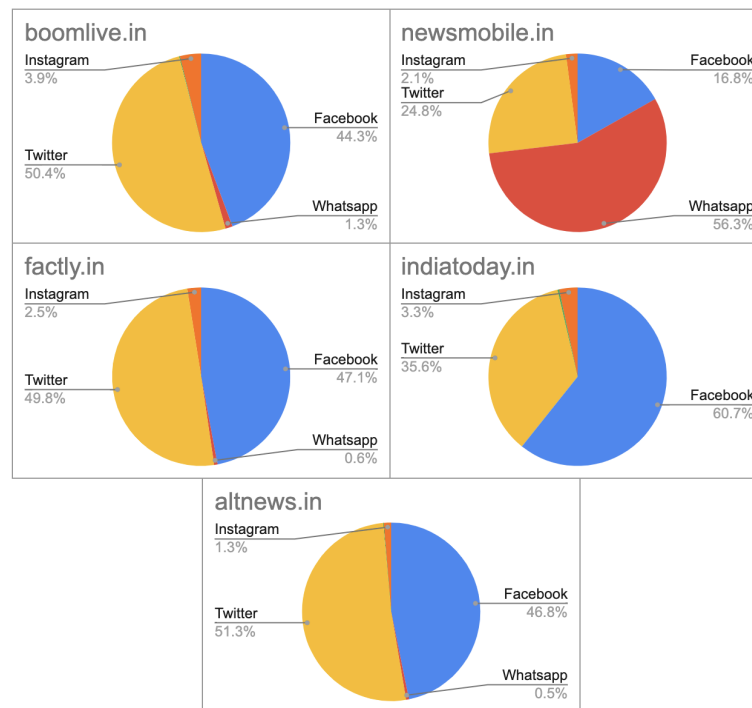


Fig. 1. Distribution of Mentions of Social Media Platforms in Published fact-Checks (Data source: Tattle archive)

5.1.1 Non-overlap of published checks.

"Plan usually is to do it before others. The newsroom is usually built to ensure that we are the first to report on a particular topic. Every morning, we make a list of all the stories that are covered by organisations and mark the ones that we are working on. We have stringent systems and do ensure that we speak to local officials and not depend on online verification." - P9

As illustrated above, certain organisations maintain a list of what other organisations are fact-checking and refer to that before deciding which articles to fact-check. There is also a tool that links fake stories to fact-checked stories, and that gives a push notification to the fact-checkers if other organisations have already done it. This highlights a lack of

469 clear incentive in attempting to check claims that have already been covered by other organisations and the prevalence
470 of a general newsroom competitiveness. This, in addition to findings from Fig 1 hint towards a potentially significant
471 disparity in general coverage of articles for checking.
472

473 The interview participants share that they use similar tenets for prioritisation of stories, especially that involving
474 the notion of impact or harm as described in the sections above. However, the archives of fact-checked articles which
475 list the stories debunked by organisations shows a large amount of disjoint. To illustrate this disparity, we show word
476 clouds of headlines in Figure 2 for articles in a 2 week sample from 1st July 2020 to 15th July 2020. The sampling is
477 done on the basis of the highest volume of fact-checking observed overall during the period of study, so that each of the
478 fact-checkers would have a high number of stories.
479

480 We observe from the word clouds that there are stark differences in the most prominent words in the headlines,
481 indicating that the selection of stories have high variation and there is a non overlap in the coverage. While all stories
482 have presence of "Vikas Dubey" and "Kerala", the rest of the words hardly have any intersection. A closer inspection
483 reveals an inclination towards stories with a religious undertone among certain organisations which is absent in some.
484
485

486 5.2 Constraints in the Verification Process

487
488 5.2.1 *Chasing the context.* Decades of history has bore witness to fact-checking evolving from objective verification
489 to advocacy-driven curbing of general misinformation. As fake news becomes more nuanced, the straightforward
490 verification processes fail to address and debunk false stories appropriately as captured by the following anecdote:
491

492 *"Hyperlocal content type of news is an emerging trend. The traceability of these kind of news is not possible,*
493 *it requires eye witnesses. For example, a woman was screaming at a child for asking another serving during*
494 *mid day meals in Andhra Pradesh." - P3*
495

496 Our interviews present an abundance of examples of stories that require significant contextual information to
497 debunk. Any checking done in the absence of primary investigation in such cases is often incomplete and vulnerable to
498 inconsistency. At the same time, timely debunking of such stories also becomes imperative to preserve context. This
499 creates a dilemma wherein organisations need to decide between allocating resources to a primary investigation at
500 the cost of a longer turnaround time, or run at the risk of losing context while optimising for the timeliness of their
501 intervention. The following quote illustrates the dilemma:
502

503
504 *"A lot of real time information will not be available in the public domain so you end up either filing an RTI*
505 *but that takes its own sweet time." - P3*
506

507 Ability to track down the context of a story is one among many resource-intensive operations organisations have to
508 undertake in order to publish well-investigated factchecks. We also expect to observe expression of inequitable distribu-
509 tion of infrastructures across fact-checking organisations in the volume of published factchecks across organisations, as
510 is supported by Fig 3. The publishing volumes of organisations remain largely consistent through every month over the
511 period of study. Clearly, organisations with sustained lower volumes of publications have had to prioritise more than
512 those with greater volumes. Unequal coverage of stories across organisations is therefore, an inevitable consequence of
513 such disparity in publishing volumes as was suggested by the previous section.
514
515

516 5.2.2 *Online vs Offline verification.* Resource and infrastructure constraints also dictate the kinds of verification
517 and investigation strategies an organisation has at its disposal. Firstly, the presence of journalistic experience in the
518 organisation is an important determiner of whether it is able to perform primary on-ground investigation to ratify
519
520

521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572



(a) AltNews



(b) BOOMLive



(c) Factly



(d) IndiaToday

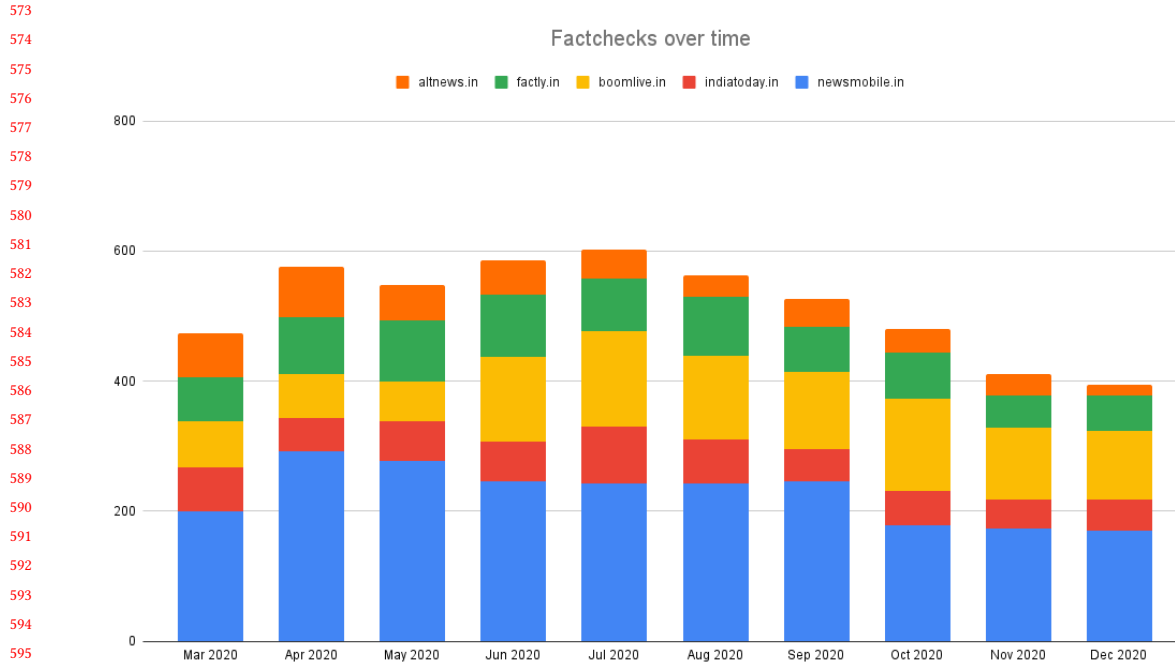


(e) NewsMobile

Fig. 2. Wordclouds from the Headlines of Published fact-Checks in July 2020

a version of events. Secondly, financial resources of an organisation dictate the volume of stories and the amount of investigative resources it can allocate at any given time period. Finally we argue that these constraints directly effect the coverage of stories across organisations, as illustrated in an evocative anecdote form P9 as follows:

"There are 2 types of stories – one, doing the reverse image search etc is more than enough. [There are many] stories where that isn't enough and we 99% times end up calling local authorities and take their word on record. There are also cases involving communal riots... There was this story on the Delhi riots where some right-wing groups had ... We sent out reporters on the ground and asked them to stand there, show me the time on his watch and show that exactly at that moment on that day and show me what he



598 Fig. 3. **Distribution of Published fact-Checks Across Organisations** (Data source: Tattle archive)

600
601
602
603
604
605

saw. Our reporter drove down and recorded the video on ground. Wherever resources permit us, if it is a place that involves taking a flight, we end up taking a call. There's nothing better or worse, there's just a healthy dose of skepticism." - P9

606 This adds a certain systemic element to unequal coverage that was previously established through Fig 3. P9's anecdote suggests that unavailability of investigating resources affect coverage of stories in a way that isn't non-random introducing a systemic bias in the kinds of stories organisations are able to verify within their infrastructure constraints. The importance of the human infrastructure as a journalist or investigator is highlighted in the following quote:

611
612
613
614
615
616
617
618
619
620
621
622

"Contrary to popular belief, fact-checking is not just about tools and techniques – the basic part still remains the same which is conventional journalism. It is not a tech kind of thing, tech certainly has a role in it, but conventional journalism like something is being said about you – so we approached you, asked you – is this correct, is there any evidence about it – concerns about any other persons, reach out to them and speak to them, speak to the people around them and ask them to provide evidence about either way. We have physical presence, who know people around and the local authorities, so we can fact-check stories from remote places which is something that many fact-checkers do not have. We ask on ground reporters to send us videos/ photos from the site and ask our reporters to send evidence that we went there. That is something we leverage occasionally" - P1

623 On the complete extreme is organisations like *AltNews* which has a total absence of on ground journalists.

625 *"One which actually needs people on the ground, we don't have people on the ground. Lets say if there is*
626 *something that is happening in the present. We do a lot of these stories and those become challenging and*
627 *sometimes we have to drop the story simply because don't have anybody on the ground to confirm." - P5*
628

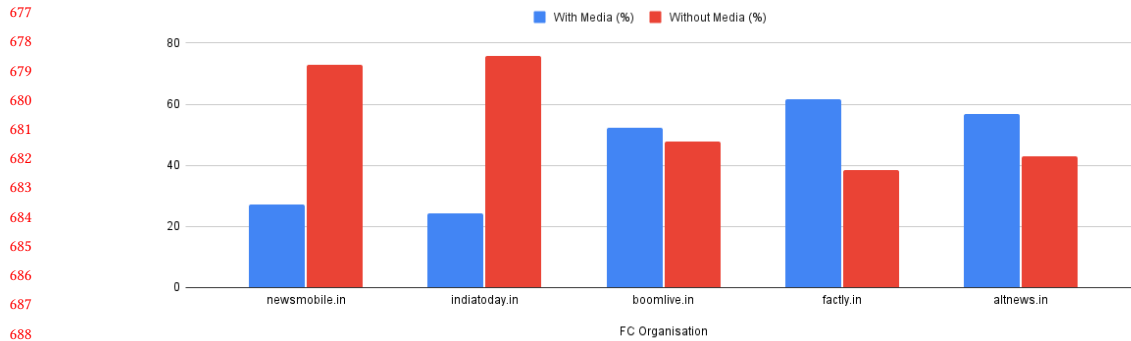
629 Due to these factors, we see an over-reliance on tools like reverse image search using tools like Yandex an Google
630 Images. The distribution of stories focusing on media related stories is also skewed. Although there are a large number
631 of opinion-based or claim-based stories that need to be fact-checked, these are harder to verify, involve a greater latency
632 in fact-checking and hence get de-prioritised amongst an influx of a large number of stories. From our archive of
633 published fact-checks, we report trends supporting the knowledge built out of the interviews. In Fig 4a, we present
634 counts of headlines containing clear evidence of the fact-check being based out of media artefacts. We search for words
635 like *video, image, photo* in the headlines to infer the substance of the fact-check. We also inspect the body in Fig 4a of
636 the published factchecks to find evidence of on-ground reporting involved by using keywords such as *visited, on-ground,*
637 *spoke to/with*. These keywords were expanded using an n-gram model to create an exhausting bag-of-words to detect
638 traces (or absence) of on-ground reporting in the article. Fig 4b clearly shows that certain organisations have a higher
639 frequency of factchecks involving media-based debunking. These also happen to be organisations limited in terms of
640 human infrastructures of journalists and on-ground reporters. On the other hand, organisations such as IndiaToday and
641 NewsMobile either have journalistic experience in their team or are known to outsource fact-checking to experts with
642 journalistic experience to allow higher counts of offline first-hand verification. NewsMobile operates largely through a
643 network of on-the-ground verifiers, which makes both their profile significantly different from the other factcheckers
644 (we also see this in Figure 5, in the high number of stories they tweet about).

649 In Fig 4a we observe that almost all organisations rely heavily on media-based debunking. However, organisations
650 with journalistic experience often seem to back up media-based verification with additional first-hand accounts and
651 offline verification. We therefore uncover yet another systemic factor that limits the nature of stories an organisation
652 is able to investigate upon. Certain stories that cannot be adequately verified by exclusively online verification are
653 therefore often out of the scope of what certain resource-constraint organisations can possibly cover.

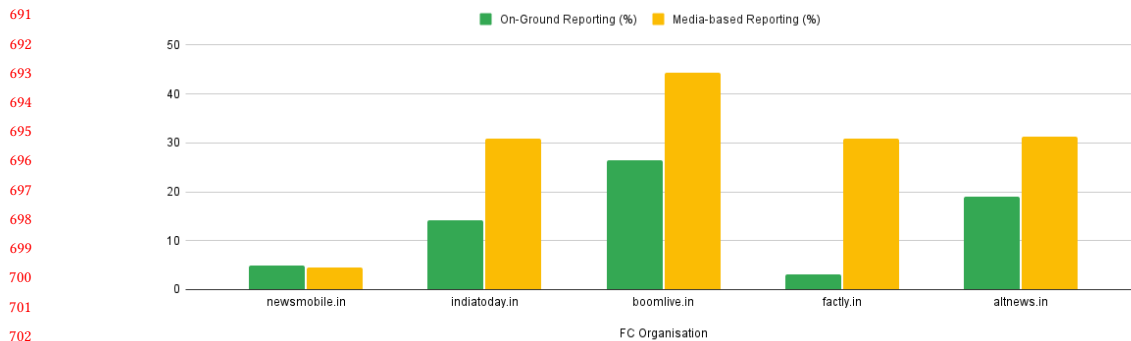
656 **5.2.3 Lack of Subject Matter Expertise and Languages.** English is not the most spoken language in India, yet most
657 fact-checkers have English as their primary language of fact-checking.

659 *"We want to do it in seven or eight languages because I feel that sometimes is important to get the right*
660 *messaging. Obviously you're constrained by funds and also that you cannot do all the languages over time.*
661 *But my plan is that I want to take it to the hinterland. Impact is greater." - P2*
662

663 In the global context, misinformation stories are often recirculated and sourced from different parts of the world.
664 When primitive online verification measures (such as reverse image search) fail, the fact-checking of stories may
665 require collaboration with fact-checkers in other parts of the countries. The IFCN networks help in this process. For
666 example, there have been instances where fact-checkers reached to fact-checkers in East Asia with misinformation
667 stories that contain number plates or road signs from their countries. It is also common for malicious stories to
668 be circulated and translated in various languages rendering the fact-checkers handicapped in terms of resources to
669 verify and/or disseminate content in non-native languages. They are also unable to write in multiple languages and
670 automated translation tools aren't in a position to enable rapid reliable translation of English language factchecks
671 to other languages. This therefore subjects the non-English audience to only a limited subset of available factchecks
672 furthering the inconsistent coverage across multi-lingual borders.
673
674
675



(a) Media-based reportage from headlines



(b) On-ground verification in factchecks

Fig. 4. Share of Media (Video/Photo) debunking based fact-Checks Among all Published Checks (Data source: Tattle archive)

"I'm a scientist...This is not what I do during the day. That's why I can invest very little time, as apposed to someone who is a full time fact-checker like [name redacted] is. Can other people have the skills to go the stuff that I am doing and do these kinds (referring to checks involving debunking of medicine-related falsities) of fact-checks? The difference in the fact-checks that we do, or I do in particular." - P4

P4's quote is a commentary on the role of subject-experts in the fact-checking ecosystem. P4's expertise in the domain of health allows her and her organisation to cover critical health-related misinformation – arguably among the most harmful kind of stories. For instance, misinformation related to alternate medicines (AYUSH Ministry) invariably requires a fact-checkers to investigate the reports submitted with the AYUSH medication. In some cases even clinical trials may need to be conducted to independently verify such claims. Thus organisations lacking the subject-matter expertise in a certain domain essentially restricts itself further with the domains it is able to monitor and investigate upon.

5.3 Publication of fact-Checks

5.3.1 *Identifying the Audience of Factchecks.* The general nature of audience who actually seek to read fact-checks are either apriori enthusiastic or have some sort of political leaning. One of our interview participants shared:

729 *"I would say anyone who is looking for discerning reader of information or consumer information, who first*
730 *of all knows that there is misinformation and now therefore trying to find the original or the corrected piece*
731 *of information is and therefore is searching." - P7*
732

733 Fake news spreads faster than fact-checked news and thus the appropriate dissemination of fact-checked news is
734 imperative to combat misinformation. When asked about how they share the fact-checked stories with the audience,
735 one of the fact-checkers said:
736

737 *"In terms of audience it all depends ones who are active, I wouldn't say it is one set of audience, we have a*
738 *WhatsApp broadcast list and people subscribe to that broadcast list. On Facebook also it works in such a*
739 *way that all those who might have shared or interacted with that content earlier will get a notification so*
740 *that is how the distribution also works. Because the content we rate on Facebook is automatically, notified*
741 *to the people who might have shared the false content earlier. that is one way of reaching new audiences.*
742 *Otherwise also we do, we run a special video series in Telegu and English. We choose most important*
743 *misinformation of the week and do short 2 minute videos on each of them. So that is one way on YouTube." -*
744 *P2*
745
746

747
748 **5.3.2 Dissemination of fact-Checks. Twitter Presence of Organisations** In this section, we analyse the outreach of the
749 various fact-checking organizations on Twitter. We first manually curate a list of 42 Twitter handles which include fact-
750 checkers and their organizations. We only include accounts that self identify as fact-checkers from a given organization.
751 It should be noted that several of these accounts are either not fact-checkers currently, or fact-checkers for different
752 organizations. However, in the time frame of this study, they were fact-checkers affiliated to the organization. We
753 collect over 142k tweets from these accounts, from which we filter out retweets, further reducing the total tweets to 93k.
754 We then analyse the social media presence of these organizations, using tweeting and retweeting patterns. Moreover,
755 we further derive insights using tweet content, followers etc. into the complex ecosystem of individual fact-checkers
756 and their affiliated organizations in India.
757

758 The distribution of the number of tweets and median retweets for each fact-checking organization is visualised in
759 Figure 5. We find that AltNews has a significant twitter presence, not only in terms of volume of tweets but also in
760 terms of engagement from their audience. Conversely, NewsMobile has the highest volume of tweets but low traction
761 in terms of retweets on Twitter.
762

763 We also find in Figure 6a and Figure 6b that AltNews has a high following and high number of verified fact-checker
764 accounts, as compared to other organizations. Furthermore, a one-way ANOVA test reveals significant differences in
765 distribution of followers (F-statistic = 6.01, p-value < 0.001 and verified accounts (F-statistic = 3.46, p-value < 0.005)
766 across all organizations. Overall, we find that AltNews has an overwhelming representation of fact-checking as well as
767 audience engagement on Twitter.
768

769 **Individual Vs Organization** In order to further understand AltNews's formidable social media presence as compared to
770 other fact-checking organizations, we dive deeper into individual fact-checkers' accounts. In Figure 7, we find that 7
771 out of the top 10 retweeted accounts are from AltNews. Moreover, we find that Mohammed Zubair and Pratik Sinha,
772 co-founders of AltNews, are significantly retweeted as compared to others.
773

774 Moreover, in Figure 8, we attempt to understand whether increased audience engagement with AltNews and their
775 fact-checkers is due to the organization's focus on narrow topics. We use Google's Universal Sentence Encoder (USE)
776 [14] to obtain high dimensional vector representations of the tweets by a user for a given event, and average out all
777 the tweet embeddings to represent a single user. The user embeddings are projected to a two dimensional space using
778
779

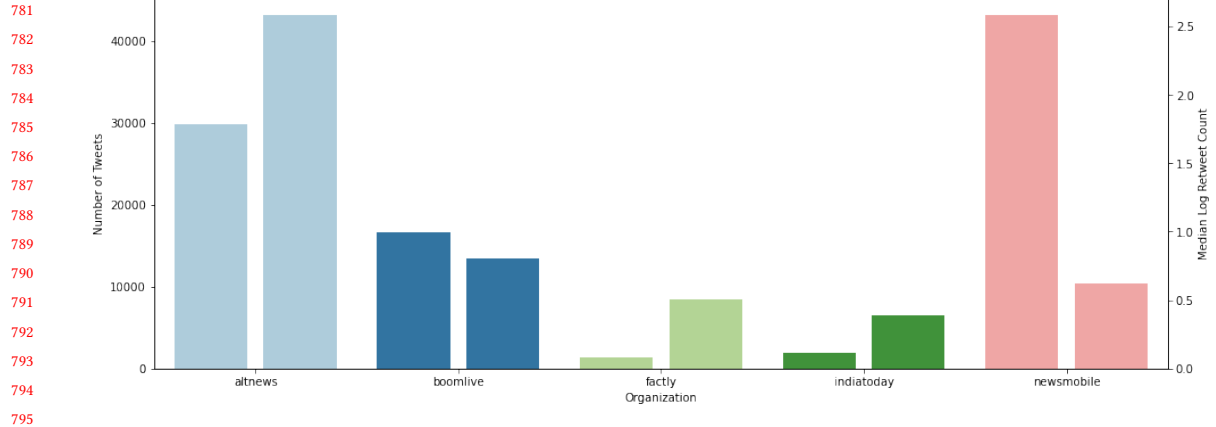


Fig. 5. Distribution of Tweeting and Retweeting Patterns Across Organizations.

UMAP [57] and clustered using hierarchical density based clustering (HDB-SCAN) [56]. We find that there is a lack of clustering of users based on their organization, implying diversity in the tweets of individual fact-checkers. Overall, we find that individual fact-checkers play a significant role in driving the success of an organization's presence on Twitter. *A Deeply Fragmented Audience on Twitter* We establish the divisiveness among the followership of the fact-checking organisations in India by visualising the followership network. We collate exhaustive lists of followers from the 5 fact-checking organisations on Twitter as a proxy estimate for their audience and readership, as each of these followers receive tweets from the organisation they follow on their Twitter feed. In Fig 9, using Gephi - an open-source software, we visualise the network of all the followers by connecting all followers of a particular organisation to a single point representing the organisation they follow. In doing so, we draw two important conclusions. Firstly, as hinted from earlier Figures 6b we see yet another evidence of the clear dominance of AltNews on Twitter even in terms of followership. Secondly, and perhaps more critically, we see a clearly divisive and fragmented audience, most of whom are loyal to a single source of factchecks on Twitter. Infact, we note that most of the pairs of fact-checking organisations in our study have less than 10% overlap in their audience with respect to the one with the larger number of followers.

6 DISCUSSION

One of the most important findings of this work is seen in Figure 9, where the true fragmentation of the consumers of fact-checked material in India is seen. While on paper, the fact-checkers do not espouse explicit ideology, they are distinct in the lack of overlap between audiences. Given that each of the three key functions of catching, corroborating, and communicating the fact-checked articles is handled differently by the three organizations, it means the average viewer may not only be excluded from a whole subset of stories that their preferred organization chooses not to engage, but also that the language and form of stories that they access form the editorializing of one fact-checker. This has the very real threat of moving consumers of fact-checked material into new polarized pockets rather than have it play the key role of verification.

833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884

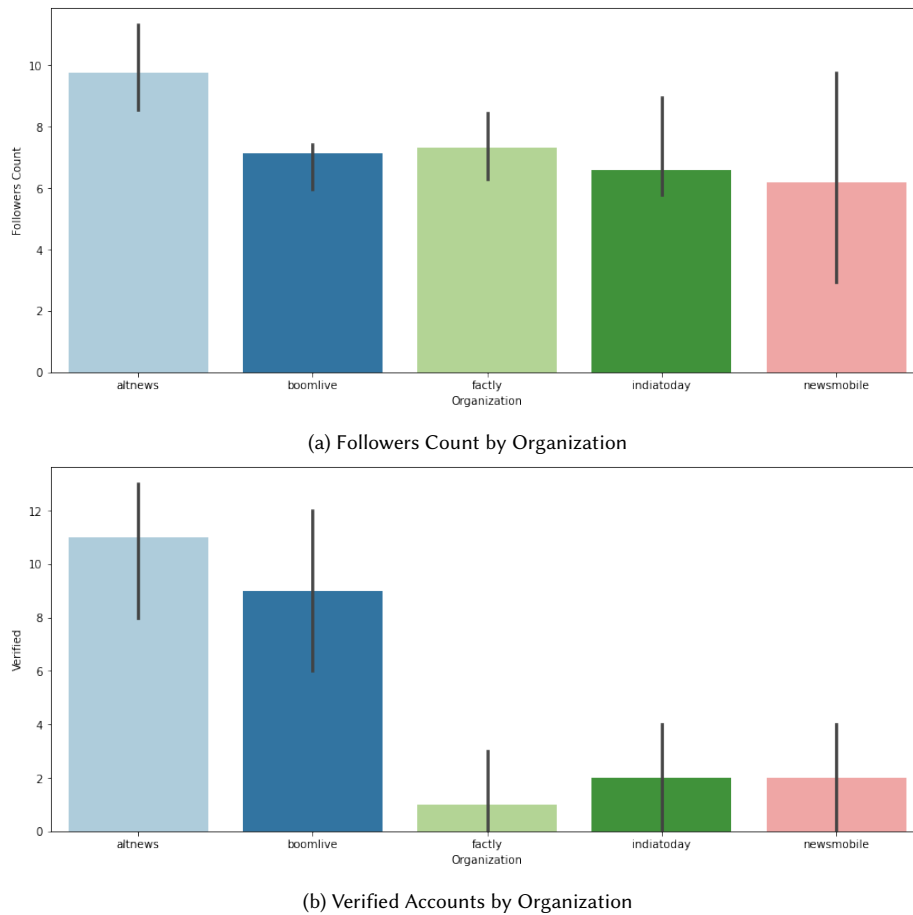


Fig. 6. **Popularity of fact-Checking Organizations in India.** Error bar represents the 95% confidence interval for each distribution. Followers count is log-scaled.

6.1 Catching

Fact-checkers report that use "virality" of the stories and potential harm as factors to prioritise the fact-checks. This model is saddled with its own problems, especially in an information environment that has come to be known for its astroturfing and manipulation of audiences. This lean towards stories that have greater affective appeal may benefit the bottom line of fact-checking organizations, but it can then highlight polarized stories more than ones that have factual problems that have potential for damage.

The virality effect also works in the opposite direction. To an extent, stories that are more viral get more eyeballs, and consequently, have greater possibility of getting debunked organically. Localized falsehoods, on the other hand, one may not even be able to effectively estimate the real impacts of.

That the "catching" process is driven by individual organizations' internal networks or mechanisms of reporting such as apps also means that they lean towards a smaller and self-reinforcing set of topics, as we see in Figure 8.

885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936

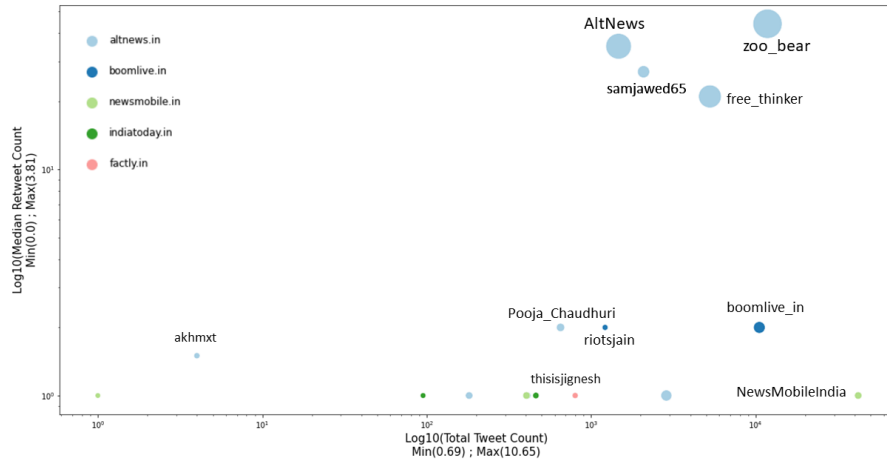


Fig. 7. Top Retweeted Individual fact-Checkers by Organization. The size of the bubble represents the number of followers.

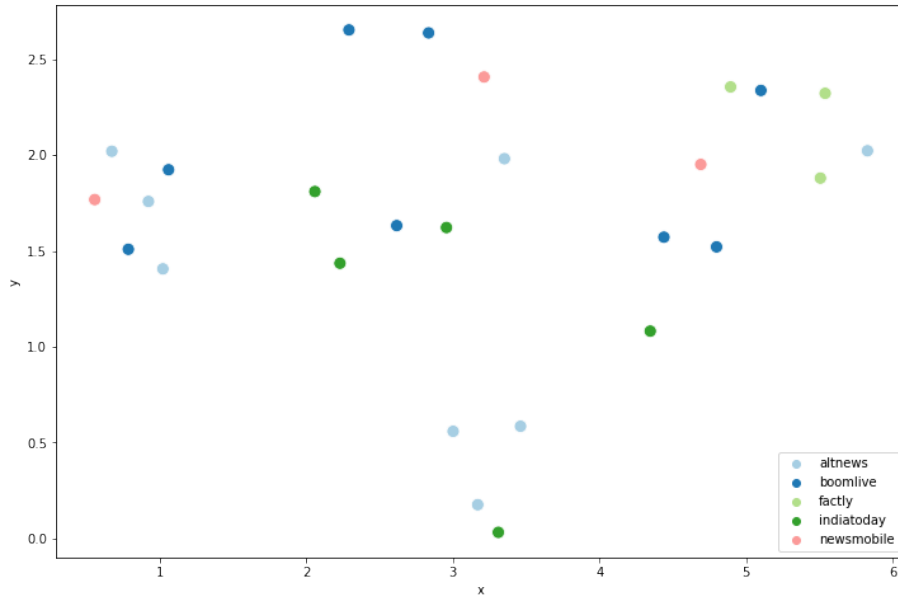


Fig. 8. Topic Overlap Across fact-Checkers and Organizations. The x and y axes represent the 2 dimensions of the UMAP projections of the tweet embeddings.

Over time, fact-checkers get a sense of what appear to be consistent patterns around stories. The vilification of minorities, the discussion of vaccines in a certain light, each of these have patterns that factcheckers come to recognize over time in their "catching" process, but this in and of itself does not reduce the frequency of these stories, as we see in the data. In other words, the affective value of these stories continue unabated, irrespective of their repeated catching and debunking.

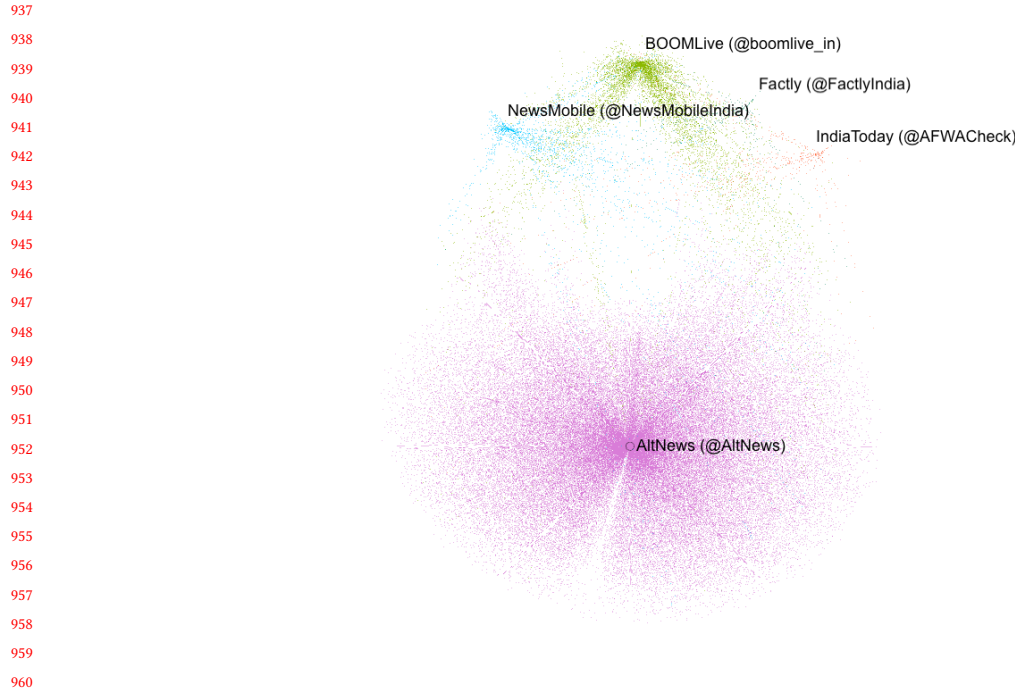


Fig. 9. **Fragmented and Highly Exclusive Followership of fact-Checking Organisations on Twitter**

A propagandist message will be spread without making direct comments but the message is conveyed. fact-checkers find their hands tied when it comes to holding repeated offenders accountable. When asked about their stance on disinformation, they admit that they cannot state "intent" of the misinformation, even when doing so would significantly improve the context and impact of the fact-checked article. A dry, run of the mill debunking which just follows the guidelines does not cut through the damage done by a highly sensitive or violent false story, which seethes through the sentiments of a community, in the very least.

This is complicated by the fact that fact-checkers repeatedly note that their challenge is not explicit lies, but innuendo, and that this innuendo is increasingly well-written, often hateful, and clearly affective enough to get viral attention. The problem then is, the "catch" has taken place in that the story is brought to the debunker's attention, but their job then is to toe a thin line between corroboration and counter-polemic.

6.2 Corroborating

While such stories show a clear ideological bias and intend to denigrate a community or an ideology, fact-checkers report that their work is increasingly driven by sticking to a formula of dry reporting. When a fact-checker corroborates a story, their point of departure needs to be whether or not they can stick a binary on the story as true or false. Furthermore, the lines laid down by organizations like the IFCN, and often the fact-checking organization's own editorial expectations is to check mainly the "facts" that need debunking rather than the intent.

As we see, there is a range of organization, each with its own style of work, with no clear attempt to talk to each other or homogenize efforts, mainly driven by the fact that "scooping" a story is clearly important given that it repeatedly

989 comes up in interviews. Indeed, some of the organizations that have long been reported for engaging in misinformation
990 have themselves come up with their own fact-checking ecologies, making clear that the battle-lines are clearly drawn,
991 and the real challenge may not be corroboration, but reach.
992

993 The active work of corroboration falls to fact-checkers on a spectrum from those who have good skills at online
994 checking, implicit knowledge and recall of issues to be able to quickly sift facts from unverifiable claims, and then turn
995 to the work of making calls. This is the major wildcard – of whether there is the need for, and consequently ability for a
996 presence on the ground. While the IFCN strives for homogeneity across its fact-checking organizations, this too is a
997 fundamentally difficult activity, since the access to resources for corroboration may vary dramatically, and the revenue
998 models for fact-checkers are simply not the same as those for mainstream journalism-centric organizations. No place is
999 this more obvious than in the minimization of hyper-local stories, and glaringly, in regional-language stories in India,
1000 which in fact reach a far higher share of Indian readership. Some of the smaller fact-checking organizations we came
1001 across in our work, which work on the smaller and more rural stories, struggle with funding and visibility, and may
1002 also lack the ability to fulfill what is needed for certification.
1003
1004
1005

1006 6.3 Communicating

1007 Finally, the dissemination of fact-checks remains an underrated challenge in the fact-checking process. We establish
1008 from our findings in Section 5.3.2, that individuals wield influence when it comes to effectively getting the audience
1009 to engage with the fact-checks on Twitter. In conjunction with Figure 9, which gives insights into the organisations'
1010 fragmented followership, we observe that individual fact-checkers who have significant outreach on Twitter in terms
1011 of followers and retweets, drive the discourse and influence which debunked stories go viral. As shown in previous
1012 works, such ecosystems, where individual accounts have disproportionate influence over the network, are susceptible
1013 to polarization, dangerous speech and manipulation [16–18].
1014

1015 The worrying aspect of this is that fact-checking organizations broadly, and some fact-checkers themselves may be
1016 expected to become institutions in and of themselves to run the specifics of an effective outreach program. As we see
1017 here, the dominance of two fact-checkers - Mohammed Zubair and Pratik Sinha of AltNews on Twitter underlines the
1018 centrality of having a solid soapbox through which to highlight a body of misinformation. But this in turn has also
1019 meant that AltNews is increasingly seen as a dissident organization, rather than one that is purely fact-checking.
1020
1021

1022 The recent detention of Mohammed Zubair under multiple charges of causing offense [64] also highlights how the
1023 lines between fact-checking and being a public voice can conflate to a heady mess for the individuals concerned. While
1024 a small number of fact-checkers may be willing to have their profiles be both very public, and subject to brutal attack
1025 from those that disagree with them, the broader question is whether the average young professional considering a
1026 fact-checking career may see this as reasonable professional endgame.
1027
1028

1029 While the focus of fact-checking organizations as well as the collectives that monitor this work are increasingly on
1030 issues of technical precision in the process of the work, it is clear from this work that the fact-checking itself is deeply
1031 dangerous work. There is need for conventional journalism to supplement debunking use of tools and technologies, but
1032 it is also needed that we rethink the role and value of strong information ecologies in our collective media environments.
1033 In particular, there is a need to consider how these matter in political and social environments that are polarized and
1034 difficult for the functioning of a free press. This indeed is also why fact-checking may be a unique blind spot of the
1035 social computing community which has obsessed over paradigms of truth that obsess over Trumpism or Brexit, whereas
1036 the functioning of media is largely free in those countries. The future of serious consideration of fact-checking must
1037 derive from work in the Global South.
1038
1039
1040

7 CONCLUSION

Gaps in the fact-checking process are also scopes for improving the fact-checking ecosystem in India. As mainstream media organisations have wide and strong networks of on-ground journalists, a greater involvement of such organisations would result in quality fact-checking for a larger volume of fake stories, and reduce the over dependence of online tools. Further, there is a need for socio-politically aware IFCN guidelines for India. IFCN lays out uniform, well generalisable guidelines to templatised the fact-checking of all stories. However, given the socio-political environment in India, the epistemology and reporting of fact-checking needs to be far more nuanced to create ripples against the fires caused by false claims. Finally, there is a pressing need for immunity and protection of fact-checkers in a politically volatile landscape. There have been several attacks against free speech of fact-checkers on social media, to the extent of their safety and integrity, when they have been involved in the process of calling out or debunking highly problematic false claims by powerful and sensitive groups. Due to this, fact-checkers need vehement protection when carrying out the crucial task of truth establishment and outreach.

8 ETHICS STATEMENT

This paper is a step towards documenting the contours of misinformation and propaganda on social media platforms, by presenting a systematic approach to identifying influential accounts in such communities. The team of researchers is comprised of urban Indians with a high degree of discerning news access and awareness, which can be a biasing factor in how the work of fact-checking is understood normatively and functionally. We identified organizations that were fact-checkers in accordance with the approval of the interviewees and the IRB process itself, but we anatomised individual interviewees. We pinpoint Twitter accounts as part of our analyses, particularly those that are highly impactful influencer accounts. However, all of these accounts are publicly available and our aggregate reporting of these accounts does not violate any Twitter regulations of privacy and consent.

ACKNOWLEDGMENTS

Hidden for blind review.

REFERENCES

- [1] Syed Ishtiaque Ahmed, Steven J Jackson, Maruf Zaber, Mehrab Bin Morshed, Md Habibullah Bin Ismail, and Sharmin Afrose. 2013. Ecologies of use and design: individual and social practices of mobile phone use within low-literate rickshawpuller communities in urban Bangladesh. In *Proceedings of the 4th Annual Symposium on Computing for Development*. 1–10.
- [2] Syeda Zainab Akbar, Ankur Sharma, Dibyendu Mishra, Ramaravind Kommiya Mothilal, Himani Negi, Sachita Nishal, Anmol Panda, and Joyojeet Pal. 2022. Devotees on an Astroturf: Media, Politics, and Outrage in the Suicide of a Popular FilmStar. In *ACM SIGCAS/SIGCHI Conference on Computing and Sustainable Societies (COMPASS)*. 453–475.
- [3] Nadeem Akhtar and Cornelius B Pratt. 2017. Pakistani government–news media relationships: how relevant are Western journalistic values? *Journalism Studies* 18, 1 (2017), 65–85.
- [4] Stuart Allan and Einar Thorsen. 2009. *Citizen journalism: Global perspectives*. Vol. 1. Peter Lang.
- [5] Ahmer Arif, John J Robinson, Stephanie A Stanek, Elodie S Fichet, Paul Townsend, Zena Worku, and Kate Starbird. 2017. A closer look at the self-correcting crowd: Examining corrections in online rumors. In *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing*. 155–168.
- [6] Arshia Arya, Soham De, Dibyendu Mishra, Gazal Shekhawat, Ankur Sharma, Anmol Panda, Faisal Lalani, Parantak Singh, Ramaravind Kommiya Mothilal, Ryna Grover, et al. 2022. DISMISS: Database of Indian Social Media Influencers on Twitter. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 16. 1201–1207.
- [7] Adam Badawy, Emilio Ferrara, and Kristina Lerman. 2018. Analyzing the digital traces of political manipulation: The 2016 Russian interference Twitter campaign. In *2018 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM)*. IEEE, 258–265.
- [8] Aasita Bali and Prathik Desai. 2019. Fake news and social media: Indian perspective. *Media Watch* 10, 3 (2019), 737–750.

- 1093 [9] Tanushree Basuroy. 2022. Fake News in India: Statistics and Facts. *Statista Dossier* (2022). <https://www.statista.com/study/68847/fake-news-in-india>
- 1094 [10] John C Bertot, Paul T Jaeger, and Justin M Grimes. 2010. Using ICTs to create a culture of transparency: E-government and social media as openness
- 1095 and anti-corruption tools for societies. *Government information quarterly* 27, 3 (2010), 264–271.
- 1096 [11] Petter Bae Brandtzaeg, Asbjørn Følstad, and Maria Ángeles Chaparro Domínguez. 2018. How journalists and social media users perceive online
- 1097 fact-checking and verification services. *Journalism practice* 12, 9 (2018), 1109–1129.
- 1098 [12] Taina Bucher and Anne Helmond. 2017. The affordances of social media platforms. (2017).
- 1099 [13] Andreu Casero-Ripollés. 2021. Influencers in the political conversation on Twitter: Identifying digital authority with big data. *Sustainability* 13, 5
- 1100 (2021), 2851.
- 1101 [14] Daniel Cer, Yinfei Yang, Sheng-yi Kong, Nan Hua, Nicole Limtiaco, Rhomni St John, Noah Constant, Mario Guajardo-Céspedes, Steve Yuan, Chris
- 1102 Tar, et al. 2018. Universal sentence encoder. *arXiv preprint arXiv:1803.11175* (2018).
- 1103 [15] Paula Chakravartty and Srirupa Roy. 2013. Media pluralism redux: Towards new frameworks of comparative media studies “beyond the West”.
 1104 *Political communication* 30, 3 (2013), 349–370.
- 1105 [16] Saloni Dash, Arshia Arya, Sukhniidh Kaur, and Joyojeet Pal. 2022. Narrative Building in Propaganda Networks on Indian Twitter. In *14th ACM Web*
 1106 *Science Conference 2022*. 239–244.
- 1107 [17] Saloni Dash, Rynaa Grover, Gazal Shekhawat, Sukhniidh Kaur, Dibyendu Mishra, and Joyojeet Pal. 2022. Insights Into Incitement: A Computational
 1108 Perspective on Dangerous Speech on Twitter in India. In *ACM SIGCAS/SIGCHI Conference on Computing and Sustainable Societies (COMPASS)*.
 1109 103–121.
- 1110 [18] Saloni Dash, Dibyendu Mishra, Gazal Shekhawat, and Joyojeet Pal. 2022. Divided We Rule: Influencer Polarization on Twitter During Political
 1111 Crises in India. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 16. 135–146.
- 1112 [19] Soham De, Anmol Panda, and Joyojeet Pal. 2022. Note: Picking Sides: The influencer-driven# HijabBan discourse on Twitter. In *ACM SIGCAS/SIGCHI*
 1113 *Conference on Computing and Sustainable Societies (COMPASS)*. 556–559.
- 1114 [20] Nicholas Diakopoulos and Michael Koliska. 2017. Algorithmic transparency in the news media. *Digital journalism* 5, 7 (2017), 809–828.
- 1115 [21] Colin Dickey. 2019. The Rise and Fall of Facts. https://www.cjr.org/special_report/rise-and-fall-of-fact-checking.php
- 1116 [22] Michael Dobbs. 2012. The rise of political fact-checking, how Reagan inspired a journalistic movement. *New America Foundation* (2012), 4–5.
- 1117 [23] Merrill Fabry. 2017. Here’s How the First Fact-Checkers Were Able to Do Their Jobs Before the Internet. [https://time.com/4858683/fact-checking-](https://time.com/4858683/fact-checking-history/)
 1118 [history/](https://time.com/4858683/fact-checking-history/)
- 1119 [24] Paulo Ferracioli, Andressa Buttore Kniess, and Francisco Paulo Jamil Marques. 2022. The Watchdog Role of Fact-Checkers in Different Media Systems.
 1120 *Digital Journalism* 10, 5 (2022), 717–737. <https://doi.org/10.1080/21670811.2021.2021377> arXiv:<https://doi.org/10.1080/21670811.2021.2021377>
- 1121 [25] Seth Flaxman, Sharad Goel, and Justin M Rao. 2013. Ideological segregation and the effects of social media on news consumption. *Available at SSRN*
 1122 2363701 (2013).
- 1123 [26] Richard Fletcher and Sora Park. 2017. The impact of trust in the news media on online news consumption and participation. *Digital journalism* 5, 10
 1124 (2017), 1281–1299.
- 1125 [27] Martin Flintham, Christian Karner, Khaled Bachour, Helen Creswick, Neha Gupta, and Stuart Moran. 2018. Falling for fake news: investigating the
 1126 consumption of news via social media. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–10.
- 1127 [28] Stephen Frantzich. 2002. Watching the watchers: The nature and content of campaign ad watches. *Harvard International Journal of Press/Politics* 7, 2
 1128 (2002), 34–57.
- 1129 [29] Ysabel Gerrard. 2018. Beyond the hashtag: Circumventing content moderation on social media. *New Media & Society* 20, 12 (2018), 4492–4511.
- 1130 [30] Tarleton Gillespie. 2018. *Custodians of the Internet: Platforms, content moderation, and the hidden decisions that shape social media*. Yale University
 1131 Press.
- 1132 [31] Lucas Graves. 2013. *Deciding what’s true: Fact-checking journalism and the new ecology of news*. Columbia University.
- 1133 [32] Lucas Graves. 2016. *Deciding what’s true: The rise of political fact-checking in American journalism*. Columbia University Press.
- 1134 [33] Lucas Graves. 2017. Anatomy of a fact check: Objective practice and the contested epistemology of fact checking. *Communication, culture & critique*
 1135 10, 3 (2017), 518–537.
- 1136 [34] Lucas Graves. 2018. Boundaries not drawn: Mapping the institutional roots of the global fact-checking movement. *Journalism studies* 19, 5 (2018),
 1137 613–631.
- 1138 [35] Lucas Graves and Michelle A Amazeen. 2019. Fact-checking as idea and practice in journalism. In *Oxford research encyclopedia of communication*.
- 1139 [36] Md Mahfuzul Haque, Mohammad Yousuf, Ahmed Shatil Alam, Pratyasha Saha, Syed Ishtiaque Ahmed, and Naeemul Hassan. 2020. Combating
 1140 Misinformation in Bangladesh: roles and responsibilities as perceived by journalists, fact-checkers, and users. *Proceedings of the ACM on Human-*
 1141 *Computer Interaction* 4, CSCW2 (2020), 1–32.
- 1142 [37] Edda Humprecht, Frank Esser, and Peter Van Aelst. 2020. Resilience to online disinformation: A framework for cross-national comparative research.
 1143 *The International Journal of Press/Politics* 25, 3 (2020), 493–516.
- 1144 [38] Faheem Hussain, Abdullah Hasan Safir, Dina Sabie, Zulkarin Jahangir, and Syed Ishtiaque Ahmed. 2020. Infrastructuring hope: Solidarity, leadership,
 negotiation, and ict among the rohingya refugees in bangladesh. In *Proceedings of the 2020 International Conference on Information and Communication*
Technologies and Development. 1–12.
- [39] Md Saiful Islam, Tonmoy Sarkar, Sazzad Hossain Khan, Abu-Hena Mostofa Kamal, SM Murshid Hasan, Alamgir Kabir, Dalia Yeasmin, Moham-
 mad Ariful Islam, Kamal Ibne Amin Chowdhury, Kazi Selim Anwar, et al. 2020. COVID-19–related infodemic and its impact on public health: A

- 1145 global social media analysis. *The American journal of tropical medicine and hygiene* 103, 4 (2020), 1621.
- 1146 [40] Joy Jenkins and Rasmus Kleis Nielsen. 2018. The digital transition of local news. (2018).
- 1147 [41] Shan Jiang and Christo Wilson. 2018. Linguistic Signals under Misinformation and Fact-Checking: Evidence from User Comments on Social Media. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 82 (nov 2018), 23 pages. <https://doi.org/10.1145/3274351>
- 1148 [42] S Mo Jones-Jang, Tara Mortensen, and Jingjing Liu. 2021. Does media literacy help identification of fake news? Information literacy helps, but other literacies don't. *American Behavioral Scientist* 65, 2 (2021), 371–388.
- 1149 [43] Beate Josephi. 2013. How much democracy does journalism need? *Journalism* 14, 4 (2013), 474–489.
- 1150 [44] Prerna Juneja. 2021. Identifying, Measuring and Contesting Algorithmically Curated Misinformation. In *Companion Publication of the 2021 Conference on Computer Supported Cooperative Work and Social Computing* (Virtual Event, USA) (CSCW '21). Association for Computing Machinery, New York, NY, USA, 256–258. <https://doi.org/10.1145/3462204.3481788>
- 1151 [45] Prerna Juneja and Tanushree Mitra. 2022. Human and technological infrastructures of fact-checking. *arXiv preprint arXiv:2205.10894* (2022).
- 1152 [46] Bente Kalsnes and Anders Olof Larsson. 2018. Understanding news sharing across social media: Detailing distribution on Facebook and Twitter. *Journalism studies* 19, 11 (2018), 1669–1688.
- 1153 [47] Shaheen Kanthawala and Jessica Maddox. 2022. Hiding in the echo chamber: fact-checking failures and individual tactics of accuracy determination on WhatsApp in India. *Asian Journal of Communication* 32, 2 (2022), 174–191.
- 1154 [48] Travis Kriplean, Caitlin Bonnar, Alan Borning, Bo Kinney, and Brian Gill. 2014. Integrating On-Demand Fact-Checking with Public Dialogue. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work and Social Computing* (Baltimore, Maryland, USA) (CSCW '14). Association for Computing Machinery, New York, NY, USA, 1188–1199. <https://doi.org/10.1145/2531602.2531677>
- 1155 [49] David MJ Lazer, Matthew A Baum, Yochai Benkler, Adam J Berinsky, Kelly M Greenhill, Filippo Menczer, Miriam J Metzger, Brendan Nyhan, Gordon Pennycook, David Rothschild, et al. 2018. The science of fake news. *Science* 359, 6380 (2018), 1094–1096.
- 1156 [50] Juliane A Lischka. 2021. Logics in social media news making: How social media editors marry the Facebook logic with journalistic standards. *Journalism* 22, 2 (2021), 430–447.
- 1157 [51] Joel Luther Mark Stencel, Erica Ryan. 2022. Fact-checkers extend their global reach with 391 outlets, but growth has slowed. <https://reporterslab.org/category/fact-checking/#article-3079>
- 1158 [52] Bertin Martens, Luis Aguiar, Estrella Gomez-Herrera, and Frank Mueller-Langer. 2018. The digital transformation of news media and the rise of disinformation and fake news. (2018).
- 1159 [53] Alice E Marwick and Rebecca Lewis. 2017. Media manipulation and disinformation online. (2017).
- 1160 [54] Donald Matheson. 2004. Weblogs and the epistemology of the news: Some trends in online journalism. *New media & society* 6, 4 (2004), 443–468.
- 1161 [55] Melinda McClure Haughey, Meena Devii Muralikumar, Cameron A Wood, and Kate Starbird. 2020. On the misinformation beat: Understanding the work of investigative journalists reporting on problematic information online. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW2 (2020), 1–22.
- 1162 [56] Leland McInnes and John Healy. 2017. Accelerated hierarchical density based clustering. In *2017 IEEE International Conference on Data Mining Workshops (ICDMW)*. IEEE, 33–42.
- 1163 [57] Leland McInnes, John Healy, and James Melville. 2020. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. arXiv:1802.03426 [stat.ML]
- 1164 [58] Paul Mena. 2019. Principles and boundaries of fact-checking: Journalists' perceptions. *Journalism practice* 13, 6 (2019), 657–672.
- 1165 [59] Nicholas Micallef, Vivienne Armacost, Nasir Memon, and Sameer Patil. 2022. True or False: Studying the Work Practices of Professional Fact-Checkers. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW1 (2022), 1–44.
- 1166 [60] Bruce Mutsvauro. 2019. Challenges Facing Development of Data Journalism in Non-Western Societies. *Digital Journalism* 7, 9 (2019), 1289–1294. <https://doi.org/10.1080/21670811.2019.1691927> arXiv:<https://doi.org/10.1080/21670811.2019.1691927>
- 1167 [61] Nic Newman. 2011. Mainstream media and the distribution of news in the age of social media. (2011).
- 1168 [62] Rasmus Kleis Nielsen, Alessio Cornia, and Antonis Kalogeropoulos. 2016. Challenges and opportunities for news media and journalism in an increasingly digital, mobile, and social media environment. *Mobile, and Social Media Environment (December 1, 2016)* (2016).
- 1169 [63] Joyojeet Pal, Priyank Chandra, and VG Vinod Vydiswaran. 2016. Twitter and the rebranding of Narendra Modi. *Economic and Political Weekly* (2016), 52–60.
- 1170 [64] Geeta Pandey. 2022. Mohammed Zubair: The Indian fact-checker arrested for a tweet. <https://www.bbc.com/news/world-asia-india-62093974>
- 1171 [65] Chang Sup Park and Barbara K Kaye. 2018. News engagement on social media and democratic citizenship: Direct and moderating roles of curatorial news use in political involvement. *Journalism & mass communication quarterly* 95, 4 (2018), 1103–1127.
- 1172 [66] Tanja Pavleska, Andrej Školkay, Bissera Zankova, Nelson Ribeiro, and Anja Bechmann. 2018. Performance analysis of fact-checking organizations and initiatives in Europe: a critical overview of online platforms fighting fake news. *Social media and convergence* 29 (2018), 1–28.
- 1173 [67] Robert G Picard. 2014. Twilight or new dawn of journalism? Evidence from the changing news ecosystem. *Journalism Practice* 8, 5 (2014), 488–498.
- 1174 [68] Ashwin Rajadesingan, Anmol Panda, and Joyojeet Pal. 2020. Leader or Party? Personalization in Twitter Political Campaigns during the 2019 Indian Elections. In *International Conference on Social Media and Society*. 174–183.
- 1175 [69] A Ross Arguedas, C Robertson, Richard Fletcher, and R Nielsen. 2022. Echo chambers, filter bubbles, and polarisation: a literature review. (2022).
- 1176 [70] Slick. 2016. Commit to transparency - sign up for the international fact-checking network's code of Principles. <https://www.ifncodeofprinciples.poynter.org/>

- 1197 [71] Dominic Spohr. 2017. Fake news and ideological polarization: Filter bubbles and selective exposure on social media. *Business information review* 34,
1198 3 (2017), 150–160.
- 1199 [72] Kate Starbird, Ahmer Arif, and Tom Wilson. 2019. Disinformation as collaborative work: Surfacing the participatory nature of strategic information
1200 operations. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–26.
- 1201 [73] Sharifa Sultana and Susan R. Fussell. 2021. Dissemination, Situated Fact-Checking, and Social Effects of Misinformation among Rural Bangladeshi
1202 Villagers During the COVID-19 Pandemic. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, Article 436 (oct 2021), 34 pages. <https://doi.org/10.1145/3479580>
- 1203 [74] Joëlle Swart, Chris Peters, and Marcel Broersma. 2018. Shedding light on the dark social: The connective role of news and journalism in social
1204 media communities. *new media & society* 20, 11 (2018), 4329–4345.
- 1205 [75] Gregory Tassej. 1991. The functions of technology infrastructure in a competitive economy. *Research Policy* 20, 4 (1991), 345–361.
- 1206 [76] Michael Trice, Liza Potts, and Rebekah Small. 2019. Values versus rules in social media communities: How platforms generate amorality on reddit
1207 and facebook. In *Digital Ethics*. Routledge, 33–50.
- 1208 [77] Peggy Valcke, Miklos Sukosd, and Robert Picard. 2015. *Media pluralism and diversity: Concepts, risks and global trends*. Springer. 1–3 pages.
- 1209 [78] MS van der Meulen and W Gudrun Reijnierse. 2020. FactCorp. A Corpus of Dutch Fact-checks and its Multiple Usages. (2020).
- 1210 [79] Feeza Vasudeva and Nicholas Barkdull. 2020. WhatsApp in India? A case study of social media related lynchings. *Social Identities* 26, 5 (2020),
1211 574–589.
- 1212 [80] Melissa Wall. 2018. *Citizen journalism: Practices, propaganda, pedagogy*. Routledge.
- 1213 [81] Wambui Wamunyu. 2021. Exploring Trust/Mistrust in Journalistic Practice: An Actor-network Analysis of a Kenyan Newsroom. *African Journalism*
1214 *Studies* 42, 4 (2021), 99–120. <https://doi.org/10.1080/23743670.2022.2026430> arXiv:<https://doi.org/10.1080/23743670.2022.2026430>
- 1215 [82] Amy Watson. 2022. Fake News Worldwide. *Statista Dossier* (2022). <https://www.statista.com/study/68847/fake-news-worldwide>
- 1216 [83] Gadi Wolfsfeld, Elad Segev, and Tamir Sheafer. 2013. Social media and the Arab Spring: Politics comes first. *The International Journal of Press/Politics*
18, 2 (2013), 115–137.
- 1217 [84] İlker Etikan, Sulaiman Abubakar Musa, and Rukayya Alkassim. 2016. Comparison of Convenience Sampling and Purposive Sampling. *American*
1218 *Journal of Theoretical and Applied Statistics* 5 (2016), 1.
- 1219
- 1220
- 1221
- 1222
- 1223
- 1224
- 1225
- 1226
- 1227
- 1228
- 1229
- 1230
- 1231
- 1232
- 1233
- 1234
- 1235
- 1236
- 1237
- 1238
- 1239
- 1240
- 1241
- 1242
- 1243
- 1244
- 1245
- 1246
- 1247
- 1248