



# Human and Technological Infrastructures of Fact-checking

PRERNA JUNEJA, University of Washington, USA

TANUSHREE MITRA, University of Washington, USA

418

Increasing demands for fact-checking have led to a growing interest in developing systems and tools to automate the fact-checking process. However, such systems are limited in practice because their system design often does not take into account how fact-checking is done in the real world and ignores the insights and needs of various stakeholder groups core to the fact-checking process. This paper unpacks the fact-checking process by revealing the infrastructures—both human and technological—that support and shape fact-checking work. We interviewed 26 participants belonging to 16 fact-checking teams and organizations with representation from 4 continents. Through these interviews, we describe the human infrastructure of fact-checking by identifying and presenting, in-depth, the roles of six primary stakeholder groups, 1) Editors, 2) External fact-checkers, 3) In-house fact-checkers, 4) Investigators and researchers, 5) Social media managers, and 6) Advocators. Our findings highlight that the fact-checking process is a collaborative effort among various stakeholder groups and associated technological and informational infrastructures. By rendering visibility to the infrastructures, we reveal how fact-checking has evolved to include both *short-term claims centric* and *long-term advocacy centric* fact-checking. Our work also identifies key social and technical needs and challenges faced by each stakeholder group. Based on our findings, we suggest that improving the quality of fact-checking requires systematic changes in the civic, informational, and technological contexts.

CCS Concepts: • **Human-centered computing** → **Collaborative and social computing theory, concepts and paradigms**; **Computer supported cooperative work**.

Additional Key Words and Phrases: fact-checking, collaboration, misinformation, human infrastructure, technological infrastructure

## ACM Reference Format:

Perna Juneja and Tanushree Mitra. 2022. Human and Technological Infrastructures of Fact-checking. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW2, Article 418 (November 2022), 36 pages. <https://doi.org/10.1145/3555143>

## 1 INTRODUCTION

“[There is] potentially an issue with.. the academic community as well, at least around fact checking.. There is loads of research done but I’m not sure [how] it makes an impact on how we do fact checking and how we report fact checking.” - *Fact-checking advocator from Full Fact* [71]

“Duke reporters lab runs *Tech & Check alerts* [56] I’d say maybe 10 times in the last few years has it helped us identify something to fact check. Lately I haven’t found it quite as useful for alerting me to something that is worth seeing.” - *Fact-checking editor from The Washington Post* [137]

With the increase in scale and diffusion of online misinformation, efforts to develop scalable technological systems for fact-checking online information have also increased. However, such systems are limited in practice [82], primarily for two reasons. First, their design is treated as a technical solution to what is often seen as a purely technological problem. But fact-checking is a complex socially-situated technical phenomena involving collaboration among multiple stakeholder

Authors’ addresses: Perna Juneja, University of Washington, Seattle, Washington, USA, [perna79@uw.edu](mailto:perna79@uw.edu); Tanushree Mitra, University of Washington, Seattle, Washington, USA, [tanmit@uw.edu](mailto:tanmit@uw.edu).



This work is licensed under a Creative Commons Attribution International 4.0 License.

© 2022 Copyright held by the owner/author(s).

2573-0142/2022/11-ART418

<https://doi.org/10.1145/3555143>

groups at various stages of the process. Yet, current automated fact-checking systems rarely take into account the insights and needs of “the human”—stakeholder groups who are central to this process. Second, automated fact-checking systems are limited in their applicability. For instance, most systems are either restricted to verifying claims about very specific public statistics by matching them against official figures (e.g., unemployment rate, inflation rate, etc.) or they are limited to identifying simple declarative claims to debunk [82]. Hence, automated fact-checking solutions fail to generalize to real-world fact-checking scenarios [82]. In other words, the rigidity of a purely technical system lacks the social flexibility necessary to support an inherent socio-technical process.

In the last decade, HCI and CSCW communities have developed a better understanding of the gap between the social and the technical [27, 66, 80, 168] and thus, are well positioned to develop an understanding of the socio-technical mechanisms underlying fact-checking. Yet, till date, we know very little about how fact-checking is done in practice and what we could do to socially and technically support the fact-checking process. In this paper, we elucidate how fact-checking is practiced by laying bare the human and technological infrastructures that facilitate and shape the fact-checking process in a fact-checking team/organization. We attempt to foreground the social by revealing the synergistic collaboration that occurs among human infrastructure—various stakeholder groups that work together to accomplish fact-checking work. We provide visibility to the stakeholder groups’ roles, needs and activities, many of which often remain invisible to the external world. We also highlight the technological infrastructure—the tools, technology, processes and policies—that supports and enables the work of the stakeholder groups. The foregrounding of the infrastructures supporting the fact-checking work helps us unravel the technical, policy and information barriers to fact-checking. Our hope is that by considering both the human and technological infrastructures underlying the fact-checking process, we might narrow the “design-reality gap” [97]—the gap that exists between the needs of stakeholder groups involved in fact-checking and design of technical systems for fact-checking. Overall, in this work we answer the following research questions.

**RQ1:** What are the various infrastructures supporting fact-checking work?

**RQ1a Human infrastructure:** Who are the various stakeholder groups involved in the fact-checking process? What roles do they play? How do they evaluate priorities and collaborate together to make decisions?

**RQ2b Technological infrastructure:** How do tools, technology and policy support stakeholder groups in performing their roles?

**RQ2 Barriers to fact-checking:** What are the various needs and challenges of stakeholder groups involved in the fact-checking process?

To answer the research questions, we adopt a multi-stakeholder approach and perform semi-structured interviews with 26 participants belonging to 16 fact-checking organizations or fact-checking teams within publication houses. We began our study by interviewing fact-checkers and editors—the stakeholder groups identified in previous works [35, 84, 154]. We discovered the existence of other stakeholder groups through these interviews and expanded our recruitment by reaching out to them using convenience [70] and snowball sampling [81]. Our participants had diverse representations from 4 continents—North America, Europe, Asia and Africa. We intentionally sampled participants widely across fact-checking teams, organizations, and countries to capture the practices and challenges emerging in this space. Our work aims at uncovering the possible human and technological infrastructures supporting the fact-checking work in teams/organizations across the regions instead of capturing the variability of fact-checking process across regions.

Our findings reveal the existence of six distinct stakeholder groups involved in the fact-checking process and the various roles performed by them. The identified stakeholder groups are: (1) *Editors*

who are responsible for overseeing the fact-checking process, including planning what topics to target and ensuring the integrity of the fact-checks produced, (2) *External fact-checkers* who are responsible for monitoring the external world (social media platforms, presidential speeches, etc.), investigating dubious claims and writing fact-checks, (3) *In-house fact-checkers* who are responsible for fixing incorrect claims present in the news stories or articles produced internally in the media/news publication house, (4) *Investigators and researchers* who conduct in-depth investigation and data analysis of persistently circulating disinformation campaigns (e.g. investigating coordinated campaigns that used anti-Ruto hashtags<sup>1</sup> on Twitter to spread misinformation<sup>2</sup>), (5) *Social media managers* who distribute fact-checks across multiple social media platforms and strategize on ways to increase audience engagement with the fact-checks, and (6) *Advocators* who spearhead initiatives to improve policies around availability of information and statistics in their countries to improve the quality of fact-checking.

By studying the roles performed by the stakeholder groups (*human infrastructure*), we establish how fact-checking has evolved from a process to debunk individual pieces of misinformation (*short-term claims centric* fact-checking) to a multi-step long-term campaign involving research, policy, and advocacy work (*long-term advocacy centric* fact-checking). We find that stakeholder groups mediate their roles via different tools (*technological infrastructure*) ranging from third party social media monitoring tools (e.g. BuzzSumo [47]), public databases, process management tools (e.g. Trello [167]), color coding schemes, to training and educational workshops. Our interviews reveal that fact-checkers are skeptical of using fully automated AI based tools. They desire algorithm explainability and involvement of humans in the decision making process as key values in the systems they would use. We also identify several technical, policy and informational challenges. For example, there are limited tools to monitor and flag content on private messaging platforms and investigate false claims in videos and content in local regional languages. We also find that in some countries, information to investigate claims from government and civic bodies is either unavailable, difficult to obtain, or is not updated periodically.

Overall, our work makes the following contributions to the CSCW/HCI community:

- Our study identifies the human infrastructure supporting the fact-checking process. We present the six key stakeholder groups involved in fact-checking and the roles performed by them. Through their roles, we show the evolution of fact-checking process from *short-term claims centric* to *long-term advocacy centric* work.
- We reveal the technological infrastructure that supports and facilitates the role of each of the stakeholder groups identified in the study.
- We offer a comprehensive understanding of complexities and peculiarities of the current practices of fact-checking and provide a holistic view of the collaborative effort among various stakeholder groups, technological infrastructure and the social-informational context in which the fact-checking takes place.
- We present the technical, policy and informational challenges that stakeholder groups face while performing their roles and state the implications of our study on the design of new tools and systems for the fact-checking process. We also highlight how in addition to the need for developing new tools and improving current platform design and affordances, there is an exigency to improve the availability and quality of data and statistics globally.

<sup>1</sup>William Ruto is the current Deputy President of the Republic of Kenya. In May 2020, several anti-Ruto hashtags (e.g. #RutoMustGo, #RutoWantedToKillUhuru, etc.) began trending on Twitter in an attempt to discredit the Deputy President.

<sup>2</sup><https://investigate.africa/opt-report-post/>

## 2 RELATED WORK

We start by presenting the definition, origin and evolution of fact-checking (Section 2.1). Next, we situate our analysis of the fact-checking process with the existing literature on human-technological infrastructure (Section 2.2) and invisible work (Section 2.3). Finally, we present the current landscape of computational research in fact-checking. We show how previous work engages with the shortcoming of the fact-checking practices and tools in a limited manner and describe how our work addresses this gap (Section 2.4).

### 2.1 Fact-checking: Definition, Origin, and Evolution

The American Press Institute defines the process of fact-checking as “re-reporting and researching the purported facts in published/recorded statements made by politicians and anyone whose words impact others’ lives and livelihoods.” [69]. One of the early examples of fact-checking emerging as an integral part of journalism was when the *Time* magazine set up a separate research department to objectively verify every printed word before releasing the publication, a phenomenon now known as ante hoc, internal, or in-house fact-checking [61]. The last decade also witnessed the emergence of post-hoc or external fact-checking which consists of publishing an evidence based analysis of claims made in any public text (e.g., news report, political speech, social media posts, etc.) after it is released to the world [85]. Today, fact-checking has emerged both as a principal part of news reporting as well as a separate entity [170]. According to Duke Reporters’ Lab, by 2019 there were around 188 active fact-checking initiatives spread across 60 countries [161]. These initiatives have incorporated a range of methodologies and data-driven journalistic practices to not only hold disinformation spreading individuals and organizations accountable through their fact-checks, but to also disseminate fact-checks in such a way that increases engagement with the public [42, 54]. With the aim of bringing together these fact-checking initiatives and in order to promote common fact-checking standards through a code of principles, the International Fact-checking Network (IFCN) was established in 2015 [138]. Major social media companies such as Facebook and Google have since then partnered with IFCN signatories to debunk false claims surfacing on their platforms [86]. While existing work describes the evolution of fact-checking from journalism to external fact-checking [83, 88], there are still gaps in understanding how fact-checking is actually practiced, the identity and role of the participating stakeholders and the various collaborations and partnerships occurring in the process. Our work deep dives into answering these missing aspects of the fact-checking phenomenon.

### 2.2 Human and Technological Infrastructures

The foundational work of studying infrastructure as a subject could be credited to Star and Ruhleder [158]. The authors consider infrastructure as “something that emerges for people in practice, connected to activities and structures” [158]. Rather than being a thing to use, Star and Ruhleder refer to infrastructure as a relational concept [158]. Scholars have since advocated for broadening the understanding of infrastructure by also including social practices, processes and flow of information [146] and have called for investigating the complexities and particularities of infrastructures in practice [48, 106, 171]. In response to this call, several research studies in Computer Supported Cooperative Work (CSCW) and related fields, such as Human Computer Interaction (HCI) have examined the infrastructures—both human and technological—supporting the diverse socio-technical systems [67, 101, 117, 130] in various contexts such as, in health-care [72, 134, 163], e-governance [51], crisis situations [120], etc. For our study, we first focus on human infrastructure which Lee et al define as “organizations and actors that must be brought into alignment in order for work to be accomplished” [117]. Scholars have used this concept of human infrastructure to denote the human

partnerships that are necessary for a successful socio-technical system [146]. Drawing on such scholarly work [67, 117, 146], we use the analytical lens of human infrastructure to “magnify the social” by rendering visibility to the stakeholder groups who collaborate to enable the fact-checking work. Highlighting the human infrastructure also allows us to focus on how collaboration and coordination is accomplished in socio-technical systems [117]. Sustaining online collaboration among various groups in an organization can be challenging [126]. Thus, within CSCW, a lot of attention has also been paid on examining collaborative [104, 113, 117] and coordination efforts [89, 119, 126, 175] in socio technical systems, determining ways to foster collaboration and coordination [78, 127, 145] and designing cooperative work tools [74, 92, 96, 107]. We complement these prior studies on by establishing fact-checking as a distributed problem that requires collaboration and coordination of the human infrastructure supporting the fact-checking process.

In a socio-technical system, human infrastructure does not exist in a vacuum [163]. It is intertwined with the technological infrastructure which is the software, hardware and processes supporting the human actors in performing their roles [140, 163]. Scholars argue that technology and human actors are mutually constituting; one mediates the other [144, 163]. Our work also borrows the concept of technological infrastructure to shed light on how use of various tools facilitates the enactment of the stakeholder groups’ roles in the fact-checking process.

### 2.3 Invisible Work of Fact-checking

Within CSCW, a lot of attention has also been paid on highlighting the invisible or the overlooked work in a process or within an organization [26, 67, 73, 147]. Invisible work can include situations where person performing the work is visible but some of the work they perform is “functionally invisible or taken for granted” [162]. Such work remains hidden in the background but is essential for collective functioning of a workplace [128]. It often includes informal work practices such as informal conversations, operational and maintenance work, etc [128, 169, 171]. There are also situations where the person performing the work itself is invisible, such as, service, design or domestic work [121, 159]. In certain complex environments (e.g. hospitals), both visible and invisible work practices can take place simultaneously [162]. In similar vein, fact-checking is a complex ecosystem that includes somewhat *visible* editorial and investigative work and *invisible* advocacy, policy and research work. Most of the prior research has looked at fact-checking as a process to debunk misinformative claims (*short-term claims centric* fact-checking) [35, 84, 85, 154]. Scholars have mostly engaged with the role of stakeholder groups such as fact-checkers and editors in supporting the fact-checking work [35, 87, 154]. We add on to the existing literature by not only expanding on the previously reported roles of fact-checkers and editors but also identifying other stakeholder groups, such as investigators and researchers, and advocates whose roles remain invisible and unexplored by prior research. We shed light on the invisible work that fact-checking organizations are doing to improve the availability and quality of information in their country (*long-term advocacy centric* fact-checking).

### 2.4 Current Landscape of Computational Research in Fact-checking

Past quantitative and computational research studies on fact-checking have primarily focused on automating multiple stages of the fact-checking process [45, 50, 79, 94, 108, 152], determining the perception and believability of fact checks [32, 75, 131] and construction of fact-check databases [112, 165, 174]. Scholars have adopted several approaches to determine the veracity of content, such as use of knowledge graphs [153], crowd-sourcing [95, 111], deep learning models [105], natural language processing techniques coupled with supervised learning techniques [94] and combination of human knowledge and AI [129]. Work in the field of multimedia forensics has also led to the development of content verification tools especially for image and video verification



such as Tineye, InVID etc. [164]. Despite the plethora of automated systems and tools available for fact-checking, our understanding of their usefulness in practise is limited. Furthermore, there is a dearth of scholarly work that engages with the limitations of the current fact-checking tools and practices ([60, 82] are a few exceptions). Our paper addresses this gap by interviewing the various stakeholder groups involved in the fact-checking process to understand the technological infrastructure supporting their work including the tools that are actually used by the stakeholders in practice, limitations of the current tools and the challenges faced by multiple stakeholder groups.

The remainder of our paper is organized as follows. We begin by describing our study method and interview protocol in Section 3. Then, we present a brief overview of *short-term claims centric* and *long-term advocacy centric* fact-checking in Section 4. Next, we answer RQ1 by describing in detail the human and technological infrastructures supporting both types of fact-checking in Section 5 and Section 6. We answer RQ2 by identifying the needs and challenges faced by the stakeholder groups in Section 7. Finally, we discuss the implications of this study in Section 8.

### 3 METHOD

To better understand how fact-checking is practised in real-world, we conducted semi-structured interviews with six stakeholder groups (N=26): (1) Editors (2) External fact-checkers (3) In-house fact-checkers (4) Investigators and researchers (5) Social media managers, and (6) Advocators. All interviews were conducted with the approval of the Institutional Review Board. We started by interviewing fact-checkers and editors employed in fact-checking organizations and publication houses. The qualitative analysis of the initial interviews (with P3, P4 and P5) as well as conversations with our contacts in the fact-checking organizations during the initial recruitment gave us a new perspective on fact-checking revealing the complex workflows that include several other stakeholder groups (apart from editors and fact-checkers) who work together to achieve the end goal of fact-checking. We then expanded our recruitment to interview people belonging to these other stakeholder groups.

#### 3.1 Participant Sampling Technique

We adopted convenience [70] and snowball sampling [81] to recruit our subjects. First, we employed convenience sampling to identify fact-checkers via Twitter search. We sent out personal recruitment messages to those Twitter users whose Twitter bio revealed them to be fact-checkers and whose accounts allowed direct messaging. The second author had established collaboration with a few fact-checking organizations. We also reached out to individuals working in these organizations. Next, we used snowball sampling for recruitment. We requested the individuals who participated in our study to further connect us with other individuals belonging to different stakeholder groups in their or other organizations. We interviewed a total 26 participants from 16 fact-checking teams/organizations including a Pulitzer Prize winning editor and journalist. Tables 1 and 2 list our participants' demographics, experience in current role, stakeholder groups that we studied, participating organizations and the continents we covered by interviewing participants based in those continents. Note that the names of the roles of various stakeholder groups were not always self-reported (with an exception of fact-checkers, news editors and copy editors), but were rather found during qualitative analysis. Therefore, these roles might not match with participants' designation in the fact-checking team or organization. For example, the professional designations of two participants performing advocacy work in their respective fact-checking organizations were *Head of Public Policy & Institutional Development* and *Partnerships manager*. Several of our participants' roles were fluid and overlapping, i.e in some fact-checking teams/organizations, at times a single person enacted several roles. For example, participant P12 performs both editorial and advocacy work. In some cases, a participant provided us with insights about more than one

P#	Gender	Exp.(yrs)	P#	Gender	Exp.(yrs)	P#	Gender	Exp.(yrs)	P#	Gender	Exp.(yrs)
P1	Female	0.67	P8	Male	2	P15	Female	2	P22	Male	3
P2	Male	0.21	P9	Female	1	P16	Female	0.5	P23	Male	1
P3	Female	2.5	P10	Male	0.42	P17	Male	0.83	P24	Male	15
P4	Male	2.25	P11	Female	2	P18	Male	10	P25	Female	17
P5	Male	3	P12	Male	4	P19	Male	4	P26	Female	1
P6	Female	0.75	P13	Female	2.5	P20	Female	7			
P7	Female	0.92	P14	Male	0.5	P21	Male	0.75			

**Table 1.** Table showing list of participants with their gender and experience (in years) in their current role. Some participants have been associated with fact-checking work for a longer duration. We only report their experience in the current role in the organization.

Stakeholder group	Organization	Continent
1) News desk editors and copy editors	Pesacheck [135], Meedan [123], First Check [115], Full Fact [71], The African Network of Centers for Investigative Reporting's Investigative Lab [23], AFP [31], Africa Check [52], The New Republic [141], The Quint [139],	Africa, Asia, Europe, North America
2) External fact-checkers	Al Jazeera [102], The Washington Post [137], DPA [6],	
3) Social media managers	Maldita [8], India Today [7], Der Spiegel [5], Fine Tip	
4) Investigators and researchers	Research & Editing, Freelance	
5) Advocators		
6) In house fact-checkers		

**Table 2.** Table showing the stakeholder groups identified in the study, the participating organizations and the continents we covered through the interviews. In the organization column, freelance refers to no association with a particular fact-checking organization/team. We aggregated the roles of stakeholders and their association with fact-checking organization/team to ensure anonymity as in some cases knowledge of network affiliation and role could potentially reveal the identities of few participants. Note that the participants that we interviewed sometimes provided insights about more than one role.

role since they either led or managed the entire fact-checking team or work(ed) closely with the other stakeholder groups and thus, were aware of the various roles involved in the fact-checking pipeline. For example, we interviewed an advocator working at Meedan, an organization that provides institutional and programmatic support to partner organizations doing fact checking work. Their job at Meedan allows them to work closely with people performing various roles at the fact-checking organizations. Thus, the participant was able to describe the responsibilities and tasks conducted by various stakeholder groups such as fact-checkers, advocators and investigators.

### 3.2 Interview Protocol and Data Analysis

All interviews were conducted between November 2020 to September, 2021. We designed a generic semi-structured interview script for our study that contained a set of broad questions about participant's and their organization's role. Based on participants' responses to these questions, we inquired them about the specific details of their roles. Thus, recruiting and interviewing different stakeholder groups did not require us to file any changes with our university's Institutional Review Board. We first asked participants to describe their role within their organization, and the function of the organization itself. We encouraged participants to share their screen and describe various aspects of their work using real-world examples. We also asked the participants to demonstrate the tools they use wherever applicable. We probed them about the role of technology in their day-to-day work and how the affordances provided by online platforms facilitates or impede their work. To get insights about how fact-checking is practised in the participant's team/organization, we asked them to describe all the steps involved in the fact-checking pipeline. We inquired about the other stakeholder groups working in their team/organization who also contribute towards the fact-checking process and how these various groups collaborate with each other. We also discussed various challenges participants face in their job. The interviews lasted between 60 to 125 minutes and

averaged over 90 minutes. All interviews were recorded via Zoom or Google Meet. The first author transcribed all the video and audio recordings. Then, two authors independently went through the transcripts and observation notes taken during the interviews and thematically analyzed them using a mixture of deductive and inductive coding scheme [46]. First, both authors conducted a deductive scan where the transcripts were coded for categories: fact-checking process, stakeholder groups, participant's role, decision making, use of tools, collaboration within stakeholder groups and challenges faced by the participant in performing their role. Then within each deductive code, inductive coding was conducted. Both authors read the transcripts multiple times to determine the codes. Then, the two authors compared and contrasted their codes with each other to refine the codes and resolve the inconsistencies. After several rounds of discussions, both authors converged on a final set of themes. We present these themes with respect to the research questions in the following sections.

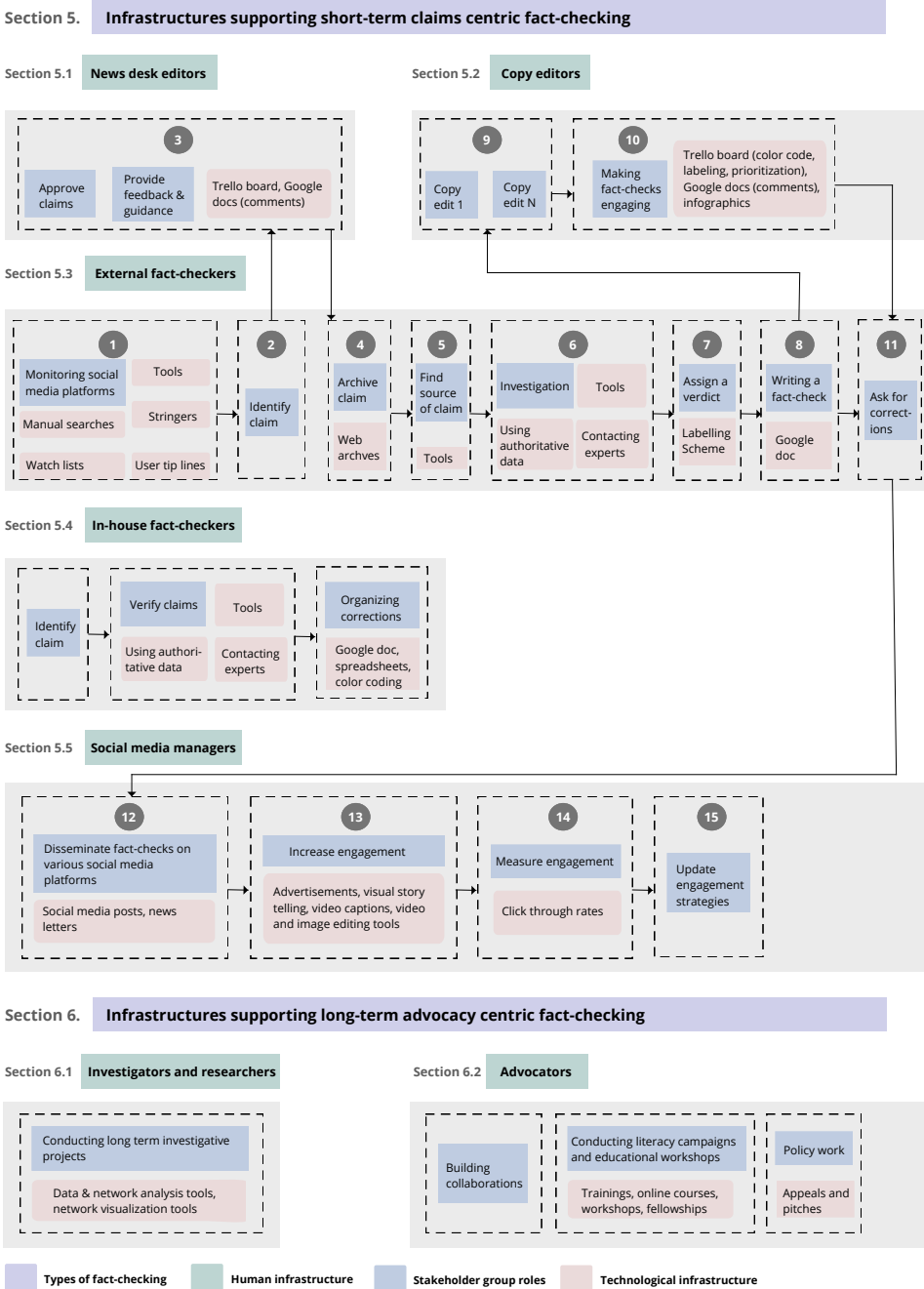
## 4 TYPES OF FACT-CHECKING: SHORT-TERM CLAIMS AND LONG-TERM ADVOCACY

Our study aims to identify the infrastructures—both human and technological—supporting the fact-checking work. We identify and present the human infrastructure by elucidating the role of six stakeholder groups that need to come in alignment to accomplish fact-checking. The six stakeholder groups are editors (news desk and copy editors), external fact-checkers, in-house fact-checkers, social media managers, investigators and researchers, and advocates. We show how these stakeholder groups' roles are supported by technological infrastructure. Through the study of the fact-checking infrastructures, we establish that fact-checking exists as both *short-term claims centric* and *long-term advocacy centric* fact-checking. In this section, we first provide an overview of the two types of fact-checking before deep diving into the infrastructures supporting them in Section 5 and Section 6. Figure 1 provides an overview of the fact-checking ecosystem including the stakeholder groups, their roles and various tools that support them in performing their roles.

### 4.1 Short-term Claims Centric Fact-checking

*Short-term claims centric* fact-checking aims at informing the public by debunking misleading claims circulating on the online platforms. It begins with fact-checkers continuously monitoring the online spaces for potentially misleading content. They identify the exact claim(s) that they want to fact-check and make a pitch to the editorial team about how they plan to debunk that claim (Section 5.3). After gaining editor's approval (Section 5.1), they archive the content, find the source of the claim, and investigate the claim by using online tools, consulting experts and employing authoritative publicly available evidence (e.g. statistics on unemployment, census data, etc.). Based on their investigation, fact-checkers assign a label indicating the veracity of the claim and then write a report declaring all the sources they gathered (Section 5.3). The written report goes through a rigorous copy editing pipeline to ensure integrity of the fact-check (Section 5.2). If the claim is false, fact-checking organizations reach out to person/organization who made the claim for correction(s) (Section 5.3). Finally, social media engagement team publishes the fact-check story/document on the social media pages of the organization and adopts several strategies to increase audience's engagement with the published fact-check (Section 5.5). Our work also examines the role of in-house fact-checkers doing *short-term claims centric* fact-checking in news and media publication houses. In house fact-checkers assist reporters/journalists in verifying and validating their articles by ensuring that the facts and quotes present in the article are correct and backed by authoritative sources (Section 5.4).





**Fig. 1.** Figure presenting the ecosystem of fact-checking, the whole or part of which could exist in a fact-checking organization or a news publication house.  indicates the two types of fact-checking (*short-term claims centric* and *long-term advocacy centric* fact-checking) introduced in the study,  presents the stakeholder groups involved in the fact-checking process (human infrastructure),  shows work done by the stakeholder groups as part of their role, and  specifies the tools stakeholders use to mediate their roles (technological infrastructure). The numbers indicate the sequence in which various roles are performed.

## 4.2 Long-term Advocacy Centric Fact-checking

“Fact checking is more than publishing fact checks. In order to change the information ecosystem., we need to spot patterns and try and do something about those patterns.. We try to influence policymakers, information producers and media to raise their standards and improve the quality of information and public debate. [Our role is] more akin to a campaigning organization. ” - P9

We find that the work of the fact-checking organizations is more than a one-off engagement with misleading claims and fact-checks (*short-term claims centric* fact-checking). Most organizations also perform *long-term advocacy centric* fact-checking. They run several investigative projects to study the misinformation ecosystem in their country (Section 6.1). They are also actively involved in advocacy and policy work where they try to influence civic bodies and policy makers to improve the quality of data, organize workshops to train organizations and journalists to do fact-checking and work towards forming coalitions among various fact-checking organizations and internet companies (Section 6.2). To demystify the fact-checking process, in the following sections, we present in detail the roles performed by the stakeholder groups involved in both types of fact-checking along with the collaborations occurring in process. For each stakeholder group, we also discuss how technological infrastructure supports the enactment of their roles.

## 5 INFRASTRUCTURES SUPPORTING SHORT-TERM CLAIMS CENTRIC FACT-CHECKING

*Short-term claims centric* fact-checking is supported by five stakeholder groups—news desk editors, copy editors, external and in-house fact-checkers, and social media managers. We present the roles played, activities performed, decisions made, and tools used by them.

### 5.1 News Desk Editors—Approving Claims and Guiding Fact-checkers

News desk editors are one of the most critical stakeholder groups supporting the *short-term claims centric* fact-checking. They decide what their team/organization is going to fact check. They approve or reject the claims pitched by the fact-checkers and guide them in their work.

**5.1.1 Approving claims to fact-check.** News desk editors are looking for newsworthy claims that impact a lot of people. First criteria of approving a claim is the popularity and reach of the person making the claim since it increases the chance of the claim spreading far and wide (“*it should be newsworthy, said by an important person,... if a popular public figure makes a claim, a lot of people are likely to be exposed to that claim, given the the bully pulpit that public figures have*”—P18). Second criteria includes number of people likely to be misled based on the context surrounding the claim. For example, a claim touching upon a communal angle is likely to impact a lot of people in a country that has “*many illiterate people..who process information based on their communal experience* (P12)”. Third, news editors approve content that is gaining a lot of traction on social media platforms by accumulating engagement in the form of likes, comments and shares. Content that has received less attention from the public is not fact-checked from the fear of amplifying the false information by inadvertently bringing attention to the false claims (“*Are people believing this or are they taking it just as a joke? Does it just have one share, meaning if I fact check it, it will just amplify the fake news and not really give the correct information*”—P11). Fourth, opinion pieces and claims that cannot be verified using sources are not fact-checked (“*we check things that have facts, that can be verified using records and information. We don’t want a situation where we see this person says this and this person says this, that will just be hearsay*”—P3). Fifth, fact-checkers and news desk editors consider several stages of harm that false information is likely to cause—physical, mental, social or emotional and prioritize the claims that are likely going to cause maximum damage to the public by affecting their health and well being.

“We have stages of harm that you have to check. Is it causing physical harm? Is it causing mental harm? Is it causing someone to lose social standing? How much effect will it have to the person.. if I don’t fact-check the story? .. We do the ones that actually have greater harm, we give them priority.. We have to get it out before most people see it to reduce the harm that it’s causing.” - P3

**5.1.2 Guiding fact-checkers.** The job of news editors does not end after approving a claim. They also guide and help fact-checkers in “*gathering sources and evidence (P11)*” for verifying stories, “*understanding concepts, finding best available data [for research],...and connecting with the experts given [editors’] long experience in the media ecosystem (P12)*”.

## 5.2 Copy Editors—Ensuring Quality of the Fact-checks

Copy editors do quality control of the fact-check story/document through multiple iterative copy editing cycles. A fact-check story or document is a report written by fact-checkers containing the claim investigated, sources used for investigation, and verdict indicating veracity of the claim. This stakeholder group acts like the *first readers* who determine whether fact-checkers have accurately interpreted the claim in the fact-check, used multiple primary sources as evidence for investigation, provided working links to the sources used for investigation and presented the evidence in such a way that it leads to a logical and correct conclusion. They also work towards making the fact-checks engaging by checking the phrasing and grammatical errors. We discuss the tasks performed by copy editors in detail below.

**5.2.1 Performing copy editing cycles.** Through our interviews we realized that a fact-check goes through two to three copy editing stages. Each stage is supervised by a different copy editor to ensure higher quality. The written fact-check is provided to the copy editor in a shared document (e.g. Google doc) where they leave comments to provide feedback. At the first copy edit stage, copy editors check the central premise of the fact-check.

“When I get a fact-check, I read through it three times to understand it before I make any change on it, before I ask any questions. I look for the claim and the debunk. Does it really hold? If there are questions about the debunk then we send it back to the news desk.. At that stage, the fact checker will pick it up and go back and try to sort out any queries that have been raised.” - P1

The second copy-edit stage focuses on refining the language and flow of the fact-check. The final copy-edit stage focuses on making the fact-check more engaging and interesting to read. We discuss this aspect briefly.

**5.2.2 Making fact-checks more engaging.** Copy editors try to keep the fact-check short, clear, crisp and interesting. They ensure that the fact-checks are written in a language that is understood by laymen. P1 spoke most candidly about the engagement aspect in the editorial process. They revealed that they often collaborate with social media managers to get feedback on the engagement aspect of the fact-checks. For example, ensuring that the country relevant to the fact-check is present “*in the title or in the blurb (P1)*” so that people could quickly determine if its of interest to them, or making certain that the verdict on the claim is placed high up in the fact-check to get more attention from the public. Copy editors along with social media managers also suggest addition of info-graphics (engaging visuals, imagery, tables, charts, etc.) to fact-checks in order to attract the eyes of the readers. The info-graphics are added to quickly communicate “*complex information in a visual manner (P2)*”. They are usually added in long-form fact-checks—the ones that debunk multiple claims and delve into the subject of the claim in greater depths.

Copy editors also ensure that the fact-check contains terms that people are most likely to search online. P1, P9, P12 and P18 talked about claim-review schema used by fact-checking organizations to allow internet companies like Google to index their fact-checks [77]. These participants believe

that using popular search terms in fact-check helps increase its visibility since search engines then rank it higher in the search results. We briefly touch upon this collaboration between fact-checking organizations and internet companies later in Section 6.2.

*5.2.3 Use of collaborative systems, labelling and color coding schemes.* In most of the organizations that we interviewed, both news desk and copy editors use Google docs to comment and provide feedback on the fact-checks. In addition, fact-checking organizations like Pesacheck, Africa Check, etc. also use an open-source project management and collaboration tool called Trello [167]. The tool provides a dashboard with a series of columns that contain cards. The columns are named so that they denote the current stage of the fact-check. For example, P2 showed the Trello board of their organization which had several stages such as *complete fact checks, copy edit 1, copy edit 2, copy edit 3, Q/A & final review, final review done, published live, and amplification done*. Trello cards are the basic functioning unit in the tool. The cards hold various information about the fact-checks including due dates, conversations, attachments etc. The tool allows editors to add colors and labels to the cards. Copy editors use labels and color coding schemes for different purposes. For example, P2 uses labels for prioritizing the fact checks (*Earlier this year there was a lot of interest in COVID, so I would label fact-check as COVID because we were prioritizing those*). P10 uses colors to indicate tasks allocated to people (*It's easier for me to see what people are working on.. If its research, it will be purple, if its documentation, it will be green.*)

### 5.3 External Fact-checkers—Monitoring, Investigating and Publishing Fact-checks

External fact-checkers are the most evident stakeholder group supporting the fact-checking process. Their role is to continuously monitor the external world for potentially false claims, investigate them for veracity, assign the verdict and publish the fact-check. We discuss these roles below. We also specify the technological infrastructures supporting the roles along with the description of the roles.

*5.3.1 Monitoring online spaces.* Monitoring content is one of the most tedious steps in the fact-checking workflow. Fact checkers monitor content reactively (in response to user tips), preemptively (before events like elections, presidential speeches, etc.) and in real time (tracking current affairs via trends and listening to conversations in real time). They monitor the content via several tools and artifacts. First, they rely on user reports and tiplines which are useful ways to access misleading content circulating in private groups and WhatsApp that are otherwise hard to access.

“We have our WhatsApp tipline and emails.. Public who wants to verify a particular piece of information send it on WhatsApp to us.. We verify those queries at our end and then send them the replies with the fact check story if possible. If we don't have fact check story, then we send people whatever information we have on the query.” - P14

Second, fact-checkers create watch lists and track social media accounts, groups, pages and websites of repeat offenders—those who posted misinformative content multiple times in the past (*“We've a database where we've tracked all accounts spreading misinformation.. We go back and check these accounts, see what they've posted on their website and personal account.”—P3*). Third, fact-checkers rely on manual searches. They follow the current events via news or Twitter trends to get updated about topics that people are talking about and track them on all online platforms. Creation of relevant search queries to track these topics is mostly a tedious *“hit-n-trial (P7)”* method. Searching for a query can give millions of results on search platforms. Therefore, fact-checkers rely on search query syntax to reduce the number of search results.

“I do not want the news content. I want user generated content. So one of the simplest tools is to write minus news (search\_query -news) so that it cancels out the major news content.” - P7

“I use keywords like “intitle” [on Google search]. [intitle:coronajihad site:facebook.com].. is showing me every search term, every post on Facebook, which has the title “coronajihad”” - P8

Fourth, fact-checkers use several tools to track content on the internet. For example, they use CrowdTangle [57] to track Facebook’s public pages and groups. Organizations that have partnered with Facebook have access to a Facebook proprietary tool colloquially known as the “Facebook Queue”. The tool aggregates potentially misleading content that is accumulating engagement on the platform. In some of the organizations, fact-checkers also rely on several third party tools (e.g. Social searcher [150], Influencer [99], BuzzSumo [47] etc.) to search and filter content on platforms since they provide them varied search filter options that the original interface of the social media platform lacks. For example, Facebook search allows to filter content by year and not by months and dates. Lastly, fact-checkers also rely on a network of stringers—“reporters who work for a publication or news agency on a part-time basis” [176]—who inform them about the misinformation “*circulating in their region..[and] language* (P11).”

**5.3.2 Making a decision to fact-check and extracting claim(s).** Once fact-checkers have identified potentially false content on the internet, they identify claim(s) in the content to verify. The fact-checkers can decide to do a short-form or a long-form fact-check depending on the number of claims they identify in the content. For their partnership with Facebook, the fact-checkers only do short-form fact-checks where they identify one claim from the body of the content. In long form fact-checks, fact-checkers will debunk multiple claims present in the content.

“The reason we do [short-forms] is to make sure we are clear, we’re not confusing the audience. This is a largely inline with the Facebook partnership where we mainly focus on one claim.” - P2

Next, fact-checkers prepare a pitch to convince news desk editors why the content needs to be fact-checked along with a plan of how they would accumulate proofs to debunk the claim. Once the claim is approved, they archive the content using online websites like archiveis [38], wayback machine [37], etc. Many times people and organizations delete the false claims made by them once they are fact-checked. Thus, archiving becomes essential to prove that the content existed.

“Once you identify the claim you archive it because what purveyors of misinformation do mostly, they delete it. So once they delete it you’re unable to read that particular claim.” - P11

After archiving, fact-checkers find the original source of the claim—“*who shared [the claim], context with which it (the claim) was originally shared* (P14)”, without which the fact-checker would not have a complete picture of the context in which the content was originally shared.

**5.3.3 Researching.** Once a claim is identified, fact-checkers collect multiple primary sources to prove or disprove the claim. They use three ways of gathering sources. First, they use quotes by experts such as doctors, physicians, meteorologists and academics. Second, they use public authoritative sources like government databases (e.g. Kenya National Bureau of Statistics), mainstream news sources (e.g. CNN), and peer reviewed scientific research papers and journals. Third, they rely on several tools. For example, fact-checkers use image and video verification tools such as InVid [100] followed by a reverse image search on search engines to collect metadata and digital trail of the image/video in order to determine its authenticity. Fact-checkers from First Check, Pesacheck, and Africa Check also reported that they depend on third party tools such as whois [177], Spoonbill [157], edit history functionality in Facebook, etc. to determine the veracity of a website or post. Through our interviews we realized the important role played by comments in fact-checking. Comments contain useful clues that help in investigating the claims.

“What we do first is read comments before checking. We found in comments that one woman said this is not the entire video, here is the link to entire video. So that’s what led us to entire video. [In comments] we see some clues, how to look for what really happened.” - P16



**5.3.4 Assigning veracity label and publishing fact-check.** After the investigation, fact-checkers assign a label to the claim that reflects its veracity. Through our interviews, we realized fact-checking organizations use a range of labelling conventions, for example, 5 point scale ranging from completely false to true, four point Pinocchio scale [136], etc. There is no commonly accepted standard for labelling misinformation. If the verdict on the fact-check is false, fact-checkers also ask the person/organization that had posted the misinformative content for correction. The last task is publishing the fact check. The fact-check contains the link to the content being debunked, description of the claim, names, quotes and links to all sources used for claim verification followed by the verdict/label. By publishing these details, the fact-checker takes the reader through the entire investigation “so that the readers can replicate [the process] themselves (P18)”.

#### 5.4 In-house Fact-checkers—Gathering Sources and Verifying Claims

In-house fact checkers are employed by publication houses to fact-check the stories produced by the journalists before they are published and disseminated. They receive the script or the news story from the journalist along with all source material that they used while researching and writing the piece. The in-house fact-checker then verifies every claim present in the story and delivers a modified story along with a list of proposed changes. Unlike the job of external fact-checkers, the job of this stakeholder group is to fix or remove incorrect claims without publicizing or calling attention to the inaccuracies [83]. The need for in-house fact-checkers arises in a publication house not only to ensure that the stories published are reporting accurate facts to readers but also to “protect the publishing house from any liability or future lawsuit (P13)”. We discuss the roles of this stakeholder group below.

**5.4.1 Identifying and verifying claims.** In-house fact-checkers verify every line present in the story. Each line usually has more than one fact to be checked, from how a proper noun is spelled, grammatical idiosyncrasies to every phrase making a claim. Journalist’s opinions and arguments, and quotes from a reputable expert are the only phrases in the content that are not verified. However, for opinions, the in-house fact-checkers check the context surrounding the argument to ensure it is “right and mainstream (P13)”. All the claims that are vague and without proper sources backing them are modified or removed from the text.

“Vague and broad facts which I. try to get people to remove.. I recently had a whole debate with someone about a line that said America is more divided today than ever. I was like, how do you check that line.. what are the sources for that kind of statement? [It] is just so broad.” - P15

In-house fact-checkers employ a myriad of techniques—top down approach, prioritization and batch processing—to identify and verify claims. The majority of the in-house fact-checkers we interviewed first perform a top down linear scan of the document to get the most central ideas “which if were incorrect, the entire piece (article) would be called into question” (P13). Second, they prioritize the claims for verification. Different fact-checkers prioritize claims differently. For example, P13, P15 and P25 work with organizations that have strict timelines for publications. Thus, they first prioritize the claims that will take maximum time to investigate. On the other hand, P20 does not have strict deadlines and prioritizes claims that they are certain are false.

“[I prioritize] a claim that I know is going to take me a while to nail down. So for example, if there’s a claim ... about someone committing a crime and I need to file a FOIA request<sup>3</sup> or go through public records or order a document from sort of government agency, I want to do that as early as possible so that I get myself as much time. You know, basically any claim that relies on other actors in order to meet to verify.” - P13

<sup>3</sup><https://www.foia.gov/how-to.html>

“When I prioritize I start with the things that I know that are wrong, and then I look at the things I think are right. And then I check all dates and all names.” - P20

P15 revealed that they “*process the claims in batches*”. Everyday they work on a batch of claims and send enquiries and suggestions regarding those claims to the journalist. This technique gives journalists ample time to respond and does not inundate them with several queries towards the very end of the schedule.

**5.4.2 Gathering sources for verification.** In-house fact-checkers use “*primary reputable sources (P15)*” to verify claims. The journalist is expected to give the primary sources that they used while doing research and writing the article. However, if the sources are missing, in-house fact-checkers look for primary documentation (like death certificate, house deed, etc.), mainstream news sources (like CNN, New York Times, etc.), academic peer reviewed journals and relevant experts to verify the claims. In case the source used by the journalist is not reliable, the usual journalistic practice is to gather a total of three to four sources to back up the claim. In-house fact-checkers heavily rely on Google search to hunt for sources. Some of the in-house fact-checkers also use Nexis search [10]—a paid service—that gives news articles, blogs, legal documents as search results. They find latter to be more effective for searching news and documents.

**5.4.3 Organizing corrections.** After identifying and verifying claims, in-house fact-checkers organize and communicate the list of questions and suggested changes to the journalist. According to in-house fact-checkers, there is no standard convention for organizing corrections. However, we found that all fact-checkers use color coding schemes for organization but in different ways. For example, P13 highlights verifiable and unverifiable claims in the Google doc with different colors and leaves comments containing information about the sources. P15 copies each claim from the doc into a separate row in a spreadsheet and then uses color coding to indicate the state of the claim (pending, verified, incorrect, etc).

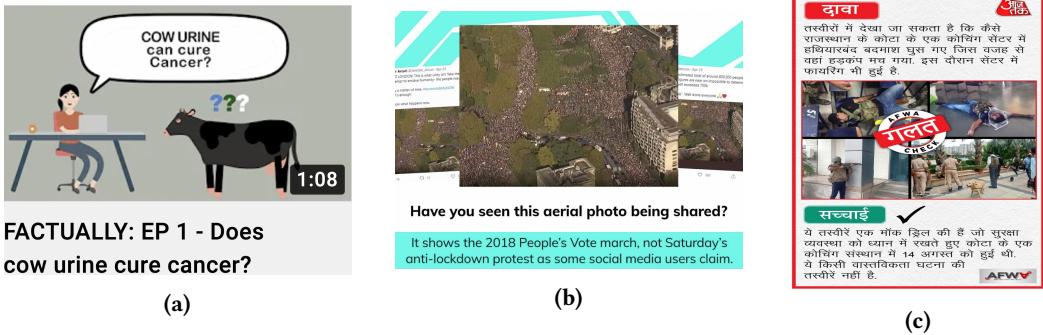
## 5.5 Social Media Managers—Disseminating Fact-checks, Increasing Engagement

Social media managers are responsible for leading all social media initiatives including publishing and disseminating fact-checks on their organization’s social media handles. They determine how to make fact-checks more appealing so that they attract people’s eyes. They also measure the engagement received by the posted fact-check(s) and based on the feedback continuously update their amplification strategies to attract more audience to engage with the fact-checks. We discuss these tasks below.

**5.5.1 Disseminating fact-checks.** Social media manager’s primary responsibility is to post fact-checks and educational tip-sheets produced by fact-checking organizations on all major social media platforms. Few organizations also have a “*WhatsApp number where they broadcast weekly newsletters containing the top fact-checks of the week (P12)*”. Social media managers want to make their fact-checks accessible to the people with disabilities. Since Twitter did not have captioning feature with its audio tweets, P17 usually posts fact-checking videos with captions. They use tools like Kapwing [19] to add captions in the videos.

“We wanted to see how can we reach, for example, blind and deaf audiences.. The audio tweets did not have captioning, they did not have accessibility features for disabled audiences. And so this is why we decided to use the video instead because videos allow us to caption and so people who can’t hear the video can still read the information.” - P17

**5.5.2 Adopting strategies to increase engagement.** Social media managers adopt several innovative ways to increase their content’s reach inorganically (via ads) and organically (via visual storytelling). We discuss some of the strategies.



**Fig. 2.** (a) A short YouTube video explaining a fact-check using comic like visuals (b) An Instagram post containing a fact-check (c) A “postcard” containing fact-check in Hindi language to be shared on mediums like WhatsApp. The single image contains the false-claim and the debunk.

- *Running advertisements:* A few fact-checking organizations that have partnered with Facebook receive free advertisement credits to promote their fact-checks on the platform. P17 delved deep into the ad usage process. They advertise fact-checks that are not bounded by time since they could be promoted via ads for an extended period. To select the audience for ad targeting, they look at three attributes, namely, the audience’s education, relevance to the fact-check and interests. They target audiences who have “*graduated from a university*” (education), live in “*countries that are most relevant to the fact-check*” being promoted (relevance) and are “*interested in news, advocacy and community issues*” (interests).
- *Visual storytelling:* Social media managers find visual storytelling to be a very effective way of getting people to engage with fact-checking content. They’ve found that many people prefer “*watching their content over reading it (P18)*”. Therefore, they convert fact-checks into a visual narrative (images or videos) before posting them on social media platforms. To create the visual content, these stakeholder groups rely on video and image editing tools. For example, social media managers in few of the organizations that we interviewed use multimedia editing tools such as Adobe illustrator [18], Photoshop [21], etc. Social media managers also create comic strips where a cartoon walks readers through the fact-checks (“*I myself had started this small comic strip thing to get more engagement.. it became quite popular and people were liking it, and sharing it*”—P7). To engage with the local non-English speaking audience, the comic strips, images and videos are also converted to regional languages.

“We translate [comics] into Swahili because there’s been this type of content, .. being done in like mainstream languages like English, French, Portuguese, But some of the more widely spoken local languages aren’t really a priority. That was a gap that we identified.” - P5

Figure 2 contains three examples of fact-checks leveraging visual storytelling techniques.

**5.5.3 Measuring engagement and updating strategies.** Measuring social media engagement is important to determine what content is gaining more traction and which engagement strategy is working. P17 informed us how they track click through rates to determine how many people are engaging with the content by clicking on the links posted by them.

“[Bitly] can tell how many people clicked on a particular link, and we can track that.. [It allowed us to] analyze how people engage with these links, how many clicks come from the newsletter, as opposed to Facebook or Twitter or Instagram? And then that allows us to know whether we need to change how we present the information.” - P17

The engagement statistics help social media teams to update their engagement strategies. For example, P17 described how they changed the number of stories published in their weekly WhatsApp newsletter from five to three after realizing that public only clicked on the top URLs.

“Earlier editions [of WhatsApp newsletter] had up to 5 stories per edition. And we were seeing that people were only clicking the top 2 or 3 links, and were ignoring the rest. So a decision was made to reduce the number of stories that we feature from 5 to 3 to make it shorter.” - P17

## 6 INFRASTRUCTURES SUPPORTING LONG-TERM ADVOCACY CENTRIC FACT-CHECKING

*Long-term advocacy centric* fact-checking aims at improving the information landscape by conducting research about various aspects of online disinformation, influencing policies surrounding availability and quality of data and statistics, conducting educational training for aspiring fact-checkers, organizing literacy campaigns for general public and forming coalitions with various fact-checking organizations and internet companies. This type of fact-checking is supported by two stakeholder groups—*investigators and researchers* who conduct long-term investigative projects that include data and network analysis, and *advocators* who are involved in policy and advocacy work. In this section, we elaborate on the role of these stakeholder groups.

### 6.1 Investigators and Researchers—Conducting In-depth Research and Investigation

Few fact-checking organizations (e.g. Full Fact, Code for Africa, etc.) have a separate team of investigators and researchers. Unlike fact-checkers who engage with individual pieces of misinformation, this stakeholder group conducts an in-depth investigation of persistently circulating misinformation and disinformation campaigns via data and network analysis.

“Part of my work at the moment is creating or developing a framework for misinformation crises. So, as opposed to individual pieces of misinformation.. looking at when responses need to go over and above the day to day..everybody who works in the kind of anti misinformation space gets together and they normally introduce new responses and new policies to manage that.” - P9

“If.. we are seeing a certain type of misinformation occurring.. every single day, debunking each individual one will not help. So what they [fact-checking team] do is now they refer that case to us. And then..[we do an] in depth investigation to try and see where is this narrative originating from.. Data analytics team is the one that does the sifting through the data sets that we obtained from social media.. Forensic team does profiling of key accounts..that we identified.” - P10

*6.1.1 Conducting long term investigative projects.* Investigators and researchers undertake several investigative projects such as verifying the backfire effect<sup>4</sup> of fact-checking (when a claim aligns with a person’s beliefs, proving that it is wrong will make them believe it more strongly), studying long term effects of conspiracy theories, examining public engagement with political news, determining how to communicate fact checks effectively, etc.

“You’ve probably heard about the backfire effect which is kind of mythical idea that fact-checking.. does more damage than good. I think the original research was repeated and the same effect wasn’t found.. We also did a project recently where we looked at.. how conspiracy theories affect people’s beliefs in the long term.. who believes and shares misinformation, how to communicate your fact checks [while] presenting them to audience.” - P9

“Projects include doing investigations into Russian disinformation or Russian influence in African countries.. investigating Chinese influence operations into .. African countries,.. human trafficking in four South African countries.” - P10

“There’s been a lot of research that was done around.. how to design a fact check to make it more engaging, how do you phrase a headline,.. how much of this information can you put into like a video. What is the ideal video length? How to get more people to interact with them and then how to try to make them much more long lasting and persistent in people’s memory.” - P5

<sup>4</sup><https://fullfact.org/blog/2019/mar/does-backfire-effect-exist/>

These stakeholder groups analyze misinformative content that is going viral on social media platforms and then trace it back to the social media accounts who started sharing it. They also conduct qualitative research by conducting surveys.

“ We have like a survey and we’ve been trying to find what kind of narratives against immigrants are more popular here in Spain. We’re trying to prepare another big survey about how fact checking works in Spain and what kind of debunk works better for us. ” - P21

The results of the investigative projects are released as dossiers (e.g. [9], [11]). The dossier provides background of the issue investigated, data collection and analysis method(s), results of investigation, and conclusion. For example, in [11], investigators study the misinformation influence operations that occurred in Uganda before the January 2021 elections. The dossier first briefly describes the political landscape in Uganda and provides examples of misinformative tweets that acted as the starting point of investigation. These tweets leveraged trending hashtags such as #StopHooliganism, #UgandaIsBleeding<sup>5</sup>, etc. to spread false narratives against the opposition party using past events from other countries. The dossier then deep dives into the methods used to collect and analyze relevant tweets and finally attributes the influence operation to the supporters of National Resistance Movement<sup>6</sup>.

Investigators and researchers use tools like python libraries for data analysis and gephi [15] for network visualization in addition to the tools used by fact-checkers.

“We were actually using an open source tool called tweepy to collect tweets from Twitter.. The network analysis is [done] using a tool called gephi [which] will be able to show relationships between tweets using the retweet function like which are the accounts that has been highly retweeted, which are the influential accounts within the network. ” - P10

## 6.2 Advocators—Influencing Policy, Building Coalitions, Conducting Educational Workshops and Literacy Campaigns

Several of our study participants revealed that fact-checking doesn’t stop with generation and distribution of fact-checks and is much more like a sustained campaign. It also includes influencing policymakers and information producers to improve the quality of information (and in turn the quality of fact-checking), building coalitions with other fact-checking organizations, social media companies, and journalistic organizations as well as providing fact-checking training to newsrooms and organizations. Such initiatives are led by advocators. This stakeholder group identifies and realizes several ways to improve the *short-term claims centric* fact checking process. They steward multiple outreach programs, policy initiatives and advocacy projects locally and globally. All the initiatives started by these stakeholder groups could be considered actions that are performed via technological and informational infrastructures such as workshops, appeals and training programs. We present the tasks performed by this stakeholder group below.

**6.2.1 Creating new generation of fact-checkers and fact-checking organizations.** The advocators conduct training, workshops and fellowship programs for people and organizations all over the world, teach them nitty-gritty details of fact-checking along with how to setup and operate a fact-checking organization of their own.

“Ethiopia.. [is] a country where press freedom is very limited, online false information often leads to offline.. And most of the media there are state controlled. So we were conducting trainings on how they can set up fact checking desks and try to be independent.” - P2

<sup>5</sup>In November 2020, Robert Kyagulanyi, a presidential candidate in Uganda was arrested on two separate occasions. The protests that broke out after his arrests were documented in Twitter posts containing hashtags such as #StopHooliganism, #UgandaIsBleeding, etc.

<sup>6</sup>[https://en.wikipedia.org/wiki/National\\_Resistance\\_Movement](https://en.wikipedia.org/wiki/National_Resistance_Movement)



“In Germany, we only got two IFCN signatories. And that’s not enough. So we try to convince traditional media outlets in the regions..to start with fact checking.. We are training them. [We have created] a community of fact checkers..and more than 600 journalists and we are doing training, encouraging them to start with fact checks.” - P19

Advocators conduct trainings through webinars or online platforms. Some organizations have set up their online learning platforms where they provide video tutorials to fact-checkers (e.g. [63]), while others have partnered with academic universities to conduct educational trainings.

“The advocacy campaigns.. on fact-checking, we conduct them virtually. We have partnerships with Kenyan institutions of higher learning, such as Daystar and Aga Khan university, which allows us to conduct webinars on fact-checking. Those sessions are attended by media professors and their students. We mainly use Jitsi, Slack and Google Meet to do the training.” - P2

“We are proud of digital e learning platform we have built at DPA. we have a lot of videos here, where we explain how to work with the Internet Archive. Those are webinars [where we] train maximum 15 people, zoom webinars means it’s live training.” - P19

**6.2.2 Pitching the importance of evidence based decision making to policymakers.** The advocates are also actively trying to reach their country’s policymakers and civil society organizations, informing them about their work and bidding the importance of facts and evidence based decision making. For example, P12 attended parliamentary researchers’ conference in Kenya where they pitched the importance of facts in informing the country’s policies and intervention programs. Similarly, P9’s organization tried to get parliamentary support to attribute electoral imprints to election campaigners during referendum and elections in the United Kingdom.

“Civil society organizations generate data that then is used by government to put in place policy intervention, say poverty eradication.. interventions in healthcare. So we just want them to understand that as you’re doing this, you also have to check your facts, don’t just rely on a news story or rely on a document or rely on a statement by a public politician to define your problem.. And so I. tell them.. their job is to tell the leaders what the data says.” - P12

“We sort of started a project, a couple of years ago about imprints. So in the UK if you distribute companion materials on paper, you have to say who is that from whereas online that’s not the case. And during the referendum and.. election.. where certain information come from online, it wasn’t attributed to campaigners in the same way that it would be offline. And so we tried to get some parliamentary support to change those rules.” - P9

**6.2.3 Organizing literacy campaigns.** The advocates actively organize literacy campaigns educating people how to critically examine information that they find on online platforms. For example, P5 shared how they partner with community radio stations, design MOOC [20] courses and frequently share tip sheets explaining people what fake news is and how they can identify it.

“We try to also work with community radio stations .. and talk through what like what fake news is,.. talking about information literacy and information disorder and how it manifests” - P5

**6.2.4 Spearheading initiatives to improve accessibility and quality of data and information.** Better decisions are made when better data is available. Quality data and information is essential because it acts as a source to verify facts in the *short-term claims centric* fact-checking. The advocates are actively working to improve the code of practices in releasing data, for example, updating it from the paper to the digital age.

“Improving the code of practice for official statistics and updating it to the internet age. And we’ve done lots of individual pieces of work trying to improve specific statistical releases because obviously they form the basis of a lot of what we do.. we try and get ..the office for statistics regulation to be a bit bolder in how they treat misuse of statistics officially.” - P9

Our interviews also revealed that data in most of the African countries is either old, or is not accessible. The advocates there are engaging with government agencies, making them aware about the problem and stressing the importance of having data publicly accessible .

“Before the census in 2019, the last census had been done in 2009. So while there are estimates available on the census data, we would use those estimates. But then the data is just not accurate when it’s estimated.. [So we] talk to the people at the National Statistics office and say we would like this data. [We] talk to people at the ministry, and trade unions telling them that this is how it would be better if you track unemployment.” - P12

**6.2.5 Building collaborations and coalitions.** Several advocates stressed on the need for all fact-checking organizations to work together, collaborate, and share resources. Such collaborations would be helpful in understanding common challenges and needs of the fact-checkers. The coalition would also better position the organizations while making certain demands from the internet companies. IFCN has played a huge part in forming such a collaboration and several advocates are actively working to expand this network.

“There’s not a lot of us working together. So that’s what I’m working on in collaboration with IFCN.. how different organizations and sectors can work better together to complement each other and collaborate and kind of information sharing and resources.., [understand] common challenges. For example, when we work with internet companies, there are certain things we might all want to ask them for which, at the moment, [only] some of us asking.” - P9

In addition to a coalition among themselves, few fact-checking organizations are also actively partnering with companies like Google and Facebook to fight fake news on their platform [55] “because they have a huge impact on the way people experience misinformation (P23)”. P9 and P12 informed us how Google is working with their organizations to make their fact-checks more visible by ranking them higher in Google search. One of our study participant (P18) was in-fact instrumental in starting the initiative.

“When we want what our fact checks to rank higher or to be more visible, so there is a back end tool claim review that we use that is integrated into WordPress.. And so the search engine looks at whatever you post as a fact check and then it ranks it higher in the matrix.” - P12

## 7 NEEDS AND CHALLENGES OF STAKEHOLDER GROUPS

In this section we answer RQ2 by presenting the challenges faced by various stakeholder groups categorized by emerging themes.

### 7.1 Skepticism Towards AI and Automation

Fact-checkers expressed skepticism and distrust towards artificial intelligence (AI) and automation of the fact-checking process. This finding resonates with prior work that revealed that the black box nature of artificial intelligence techniques and machine learning algorithms make their inner workings unintelligible to humans thereby decreasing users’ trust in their outputs [133, 149]. P4 divulged that they are skeptical of Facebook’s AI-based tool that aggregates potentially misinformative content for fact-checkers to verify and rate on its platform because they think that “*algorithms hide a lot of stuff*”. They believe that fact-checkers or independent organizations should be responsible for aggregating content to fact-check on the social media platforms rather than “[*companies*] that are running the platforms.”

Fact-checkers understand that machine learning models work when there are lots of similar data for training and pattern recognition. Thus, P18 doubts AI’s capability to detect falsehood in politicians’ statements which can be very diverse in language and topics. They also doubt AI’s capability to differentiate between a true and a false statement especially when there are only

subtle differences between the two. P19 does not believe that fact-checking could be automated since its a complicated process that requires several humans to discuss and make decisions.

“ I think [AI based tools are] going to be less useful for most.. politicians, because the problem is people don't repeat stuff the same way and the addition of a word or two can make a huge difference.. I don't think a computer is ever going to be able to figure that out. ” - P18

“[Fact-checking is] such a complex process, for example, extracting a claim,- what's the underlying meaning of a certain claim, how to understand it. Even [for humans] it's a process of discussing and then deciding. So I think it still will be humans work in a way. ” - P19

Despite the skepticism towards AI, we found that some fact-checking organizations have indeed adopted AI based tools and a few others showed a willingness to adopt such tools. However, such tools are only acceptable for low stake scenarios of monitoring content on social media as compared to high stake scenarios of assigning a veracity label to the content. This observation is inline with recent work that found that the use of AI is more acceptable in low stake compared to high stake decision making processes [41, 148].

“I will not trust any algorithm or any AI to flag something as right or wrong, at least not at this stage.. I prefer AI only to give us a curated list, flag us that this is something that we should look in. Make a tool that picks up the signals that [indicate the content is] misleading and makes a curated list. ” - P22

“If you're serious with fact-checking, you cannot replace it by automatization or things like that. But it could be helpful in [social-media] monitoring. ” - P19

To increase fact-checkers acceptance of AI based tools, P16 and P12 stressed how its essential to have humans in the fact-checking loop. P12 further said how they would only trust the AI output if the tool is able to explain how it arrived at a particular conclusion. Past studies have also developed human-in the loop AI systems [155, 178] and tried to make them more explainable in order to foster trust in them [24, 58, 118, 125, 149].

“I think at a certain point, AI stops and human being needs to come in to verify.. I don't think that AI can really do exactly the same what we can do. Not yet at least.” - P16

“ The manual process would still have [to be there]. I would be willing to use it [automated tool] to see how it arrives at that conclusion. So if you say this is misinformation, and these are the sources of data that we're using to make that. And we check and find that the algorithm.. is not using them out of context. So at that point, we would be in a position to say let's check it. ” - P12

## 7.2 Need For Tools and Limiting Social Media Affordances

We found stakeholder groups divulging several challenges related to the tools they used and the affordances provided by social media platforms. In the process, we also discovered a few needs with respect to tools and systems. We present them next.

*7.2.1 Monitoring Social Media Platforms is Manual, Time Consuming and Difficult.* Fact checkers complained about the information overload problem. With the emergence of a variety of social media platforms, the amount of online information is increasing exponentially. However, for the fact-checkers, searching for misleading content mostly remains a manual task. In addition, generation of search queries that could lead to potentially dubious content is still based on hit and trial method. As discussed in Section 5.3, only Facebook has provided a tool to its partner fact-checking organizations that can aggregate potential misinformation. While a number of social media monitoring tools have been developed to identify and aggregate misinformation, particularly on Twitter (such as [34, 50, 98, 142, 152, 166], etc.), other social media platforms (e.g. YouTube, Google, Yandex, etc.) where misinformation is equally prolific also need attention.

“The search is not easy to be honest.. it takes a lot of time to actually find the content because we only have a manual method to do that, hit n trial method to do that.. I have even scrolled to the point where YouTube shows me no more results. So that’s how manual it gets.” - P7

“If there’s some.. tool that helps filter..different types of misinformation will help because now you have information overload and you don’t know what to choose and what not to choose.” - P3

**7.2.2 Limitations of platform affordances.** Many fact-checkers complained how online social media platforms’ affordances become a hindrance for searching and filtering content. For example, inability to search for posts and comments on Facebook, lack of search trends feature on platforms (with an exception of Twitter), unavailability of fine grained search filters, inability to download content on platforms like Instagram and inability to search for same videos that were uploaded with different keywords (title, description, etc.) on YouTube are some of the limitations of platform affordances.

“Facebook is actually tricky to be honest, because there’s no one way to search content. you can only search people..pages and groups, etc.. but I need user generated content.” - P7

“Youtube search engine is not very good. It will just show you only videos which are popular which has no use. Misleading videos won’t have too much views, but they have a lot of uploads. What happens is that since a lot of people.. upload videos using different caption, different keywords. So.. [there might be] 10 versions of the same thing.” - P8

**7.2.3 Systems and tools needed to detect misleading claims on private message platforms.** Privacy settings on social media groups (e.g. Facebook) and end-to-end encryption in messaging platforms (e.g. WhatsApp) act as a hindrance in accessing misleading content circulating on these platforms. The fact-checkers informed that they need tools that allow them to access and flag content on these platforms. Recent research work has focused on building crowd sourced WhatsApp tip lines for discovering content to fact-check [109, 124]. However, being able to track or report private messages via tools necessitates serious ethical considerations. Private messaging platforms give users a false sense of security, thereby making them share sensitive information (e.g. clinical records of patients [122]), without anonymizing it [91] and hence, expose users to privacy risks [43].

**7.2.4 Overload of tools.** During our interviews, fact-checkers showed and talked about several tools that they use during fact-checking. One of the participants P8 showed us around 15 tools. There is a tool overload problem. P3 suggested building a single tool that could provide all functionalities that fact-checkers need. However, since fact-checking is a complex task involving multiple steps, it is difficult for a single tool to cater to a divergent set of requirements and functionalities [166]. Thus, building a suite of purpose-built tools that cater to specific needs of fact-checkers in the various steps involved in fact-checking would be more useful for the fact-checkers.

“[If one could] put all these tools in one, .. So I don’t have to look for different tools when I’m analyzing a video or an audio. I can do it in one place instead. Like.. I can use 10 tools to analyze the video. But if we can have one..[tool that] gives you the information you need.” - P3

**7.2.5 Need for specific tools.** Despite the problem with tools overload, existing tools lack task-specific functionalities. For instance, a lot of steps involved in video fact-checking are manual. While tools like Invid and reverse image search are available to verify a video, they are only useful to check if the video is digitally altered or used in a different context than the original. For all other videos (e.g. videos with conspiracy theories), claim extraction and claim verification process is manual. For lengthier videos, this process could get even more tedious.

“There is no tool to debunk [conspiracy theory videos]. For example, I cannot do a reverse image search, I cannot divide the video into key frames, because it is a narrative that is false.. We have to really go line by line and see what the person is saying, and then all we can do is search quotes from different organizations to [verify the claim].” - P7

Fact-checkers believe that “*efforts against misinformation advance much more and much quicker in English*” (P23) than in other regional languages. They need tools to transcribe videos in local regional languages.

“The problem with transcribing videos online or using any other software is languages because India has so many languages and we tend to get a video and information in every possible language. So it becomes difficult to have one dedicated tool to transcribe all our videos.” - P14

Editors revealed that editorial work is mostly manual. P1 spoke about how editing work is “*still very human*” and the human resource management tools like Trello could be further strengthened. For example, currently, “*there are a lot of issues of accountability*” with the Trello board, it lacks control features because of which “*anybody can move a card anywhere*.”

**7.2.6 Getting organic engagement for fact-checks without advertisements.** P17 revealed that while on some platforms (e.g. Twitter) it is easier to “*get organic reach and engagement*” by adopting appropriate strategies, it is “*a big challenge to get [same].. attention without advertising*” on other platforms (e.g. Facebook).

### 7.3 Issues around policy and information infrastructure

Stakeholder groups discussed several challenges surrounding information availability and quality. They understand that quality data is essential not only for developing AI based automated tools but also for investigating claims.

**7.3.1 Need to improve information quality before automation.** AI models are as good as the data available. If the information against which the claims are to be verified is missing or of low quality, the models would never work.

“Automation is tricky because.. in a place like East Africa ..information is not readily available. You cannot say that I’ll go to this site and get this information so that when these numbers are presented it can easily be automated.” - P12

P9 raised an interesting point along the same lines. They explained how success of automated fact-checking is dependent on the accessibility and format of statistics and information available. The data has to be in the same format for the machine to be able to understand it. Their organization is working with institutes around the world to improve statistics globally.

“[We want] statistics being published in a kind of open accessible and consistent format. So, for example, like any symbols that are used to show a caveat about data needs to be the same so that our machine can understand them each time.. My colleague.. is working with the open data Institute in UK and also globally.. to [determine] whether there are ways of improving statistics so that automated fact-checking can work.” - P9

P22 elicited how data in their country is not available in a user friendly or machine usable format.

“The data, you will find that it is in a PDF.. file or a photograph on the website..then how do we use it?...You need the data ..[to be] put into a excel sheet so you can clean it. Data should be in a format which is user friendly so it can be used, [it should be] machine learning friendly so that the machine can pick up and use the data. ” - P22

The aforementioned concerns elicited by fact-checkers are in line with the prior research that has also listed absence of structured and quality data as well as lack of adherence to data standards, as few of the major challenges faced in the field of big data [30, 59]. Lack of harmonization between data sets makes data integration a complicated and time consuming task [1, 2]. Data integration is necessary in various scenarios related to fact-checking [116], for example, querying different databases originating from different sources to determine the veracity of content or determining whether a piece of content has already been fact-checked by searching in various fact-checking



databases [49, 116]. Thus, scholars have stressed on the need to adhere to common data standards to help facilitate data integration and reuse [44]. There have been a number of cross country initiatives to set common standards for data in various domains. For example, in 2017, European Medicines Agency held a meeting to discuss the opportunities and challenges in applying a common data model to healthcare data across the countries in Europe to support regulatory decision-making [3, 4]. Furthermore, several advocators (Section 6.2) have also been spearheading initiatives to improve the data quality in their respective countries.

**7.3.2 Lack of information sources.** Our interviews with fact-checkers in the Global South revealed that the information needed to investigate claims either does not exist or is not updated periodically.

“There’s a lot that we don’t cover because of lack of sources.” - P2

“In Kenya,.. demographic Health Survey, which .. shows the health situation in the country, the last one was done in 2014, this is 2020. It’s just too old, and we can’t use it.” - P12

**7.3.3 Difficulty in getting information for research from civic organizations.** In most cases, fact checkers need multiple sources to debunk a claim. To obtain these resources they rely on public data sets and information from government and civic organizations. Six fact-checkers working in African countries and the Balkan regions informed us how information needed for research is not publicly available and getting it from officials is a long and difficult process.

“It’s so hard to get information from the government.. because everybody wants to protect themselves. They don’t want to give you the information that you actually need.” - P3

**7.3.4 Algorithmic bias against local content.** P5 complained about algorithmic bias in terms of how content indexed in search engines is skewed towards the Global North region making it very difficult to search for local content in regional languages of the Global South.

“The way the algorithms works was very I’d say Euro-centric or like North America [centric].. Trying to find tools that would enable me to find.. stuff that’s not in English, like things in local languages was quite challenging.. it would take a while for some of the stuff from from this side to be indexed on.. Google searches and other platforms .. So, there’s sort of algorithmic bias when it comes to.. find stuff like that.” - P5

## 7.4 Emotional cost of fact-checking

In addition to the manual labor involved in fact-checking, fact-checkers also face significant emotional toll and stress in their job. They are often victims of online threats and abuse from users and conspiracy theorists whose posts they were tasked to debunk. Manually scanning through misinformative content about certain topics, such as riots, conspiracy theories, also has adverse affects on their mental health.

“It becomes kind of very stressful job. Seeing all violence and getting into each and every detail, kind of takes a toll on your mental health. And then again you have to listen to abuses. And the situation is worse when you like put this content online and then people attack you” - P8

“We are exposed to threats, these conspiracy theorists are very aggressive and I worked only for like a couple weeks when I saw a CrowdTangle post with my own photo saying, this is your censor and it was very unpleasant feeling to find that.” - P16

While previous work has studied emotional labor and psychological symptomatology in content moderation work [39, 62, 110, 143, 160], no study has investigated the emotional cost of fact-checking work. Studying human costs underlying the fact-checking process and determining wellness interventions to psychological effects of fact-checking are another fruitful avenues for future research.

## 8 DISCUSSION

In this section, we discuss how our study renders visibility to the human and technological infrastructures supporting the fact-checking work and the collaborative efforts involved in the process. We also discuss the needs of the stakeholder groups and the implications of our findings on the future research directions in fact-checking.

### 8.1 Rendering visibility to the human infrastructure of fact-checking

Till date the primary objective of fact-checking is considered as debunking misleading claims. Based on this objective, prior work suggests that fact-checking can only influence three constituencies—people, journalists and political operatives [35]. By identifying and examining the human infrastructure—the stakeholder groups that need to be brought into alignment to accomplish fact-checking, our work provides a means to think about fact-checking as a multidimensional initiative which then assists in understanding 1) the invisible aspects of the process, and 2) the other ways fact-checking process can have an influence. Our work shows that fact-checking is supported by several processes, such as editorial work, social media engagement work, in-depth research and data analysis, as well as advocacy and policy work that might not be visible to the external world. Through the study of these processes, we establish how fact-checking has evolved to include both *short-term claims centric* and *long-term advocacy centric* fact-checking. We make visible the efforts that fact-checking advocates are putting in to improve the availability, accessibility and quality of data and statistics by aligning the focus and interests of governments and internet companies with fact-checking organizations. Through these efforts, the fact-checking organizations are not only improving the information landscape of their country but in turn are also improving the quality of (*short-term claims centric*) fact-checking itself. Rendering visibility to the work of the human infrastructure and the invisible processes of fact-checking has helped in uncovering the needs—both social and technical—of the entire fact-checking ecosystem consisting of all the stakeholder groups. The knowledge of needs of the ecosystem could enable design and development of tools and policies to support various aspects of the fact-checking process.

### 8.2 Collaborative efforts in the fact-checking process

Our study highlights fact-checking as a distributed problem where collaboration takes place at multiple stages among people with different skill sets within and outside the fact-checking team/organization. First, collaboration occurs among the stakeholder groups: editors, fact-checkers, social media managers, researchers & investigators, and advocates (refer Figure 1 for an overview). Second, collaboration extends to the outside world with experts such as doctors, oncologists, academics, etc. whose expertise is needed to investigate dubious claims. Third, fact-checkers collaborate with civic and government organizations during the investigative stage to access data and statistics related to the claims under investigation. Fourth, in parallel, advocates reach out to policymakers to influence policy by highlighting the challenges faced by fact-checkers. Fifth, several internet companies, like Facebook collaborate with fact-checking organizations to fact-check and debunk misleading claims on their platform. Finally, collaboration occurs between social media users and the stakeholder groups at two stages: 1) at the content monitoring stage where users report dubious claims that they encountered online directly to fact-checkers via tip-lines, and 2) when users engage with fact-checks disseminated by social media managers. Rendering visibility to the collaborative efforts in the fact-checking process has multiple benefits. We discuss a few.

*8.2.1 Increase in efforts to foster collaborations.* Making visible the collaborative efforts in the fact-checking ecosystem can lead to efforts and policies that foster these collaborations. There has been growing research on how to make users engage with the published fact-checks [36, 53, 68, 76, 179].

For example, fact-checking organization Pesacheck created a Twitter bot named *debunk bot* that detects tweets containing URL(s) to misinformative content and replies to them with the link to the fact-check that their organization has published [13]. Similar investigations and efforts can be put to support other collaborations, such as between experts and fact-checkers. There has been only a handful of recent efforts in this direction, for example, Meedan’s Digital Health Lab<sup>7</sup> and Facebook’s Journalism project<sup>8</sup> support fact-checkers in debunking health related misinformation by connecting them with health experts and providing them with resources on the health topics that they are covering. Imagine a private social media platform consisting of separate communities (like subreddits) of fact-checkers, and experts from different fields (from doctors, journalists, meteorologists, to university librarians and professors) to facilitate easy, and targeted communication and information sharing. Fact-checkers can post questions in relevant communities, seek quotes from experts, and get suggestions for online and offline resources to support their investigations. Fact-checkers can also share with each other their concerns or information about new tools that they discovered. Such a platform could facilitate fact-checking organizations in addressing online misinformation effectively and in a timely manner.

**8.2.2 Revealing the value of fact-checking work to internet companies.** In recent times several platforms such as Google search, YouTube, and Google images have started actively using fact-checks produced by fact-checking organizations with their search results to help determine their validity and truthfulness [12, 16, 22]. Our study also revealed how fact-checking organizations are collaborating with internet companies and allowing them to index their fact-checks. This finding contributes towards the HCI scholars call of making people aware of “the value their data brings to intelligent technologies” [17, 173]. Our work highlights the value that fact-checking is bringing to the social media companies who regularly use the fact-checks to regulate the content on their platforms and provide reliable information to the users. Previous scholarly works have raised questions on whether volunteer created content such as Wikipedia articles should receive more economic benefits from the internet companies [172, 173]. On similar lines, we want to raise the question whether fact-checkers and fact-checking organizations should also receive more economic benefits for their work.

**8.2.3 Revealing power dynamics in collaborations.** Our study also sheds light on the power dynamics of collaborations happening in the fact-checking ecosystem. For example, our work shows how editors have the power over fact-checkers in determining what kinds of misinformative claims should be prioritized. Additionally, in certain situations, fact-checkers depend on civic and government organizations to get data for investigation. Tools released by social media companies, such as Facebook Queue (refer Section 5.3.1) also dictate what kind of claims fact-checkers investigate. These companies also have the power to remove the fact-checks from their platform. For example, Facebook removed a fact-check on abortion after receiving complaints from Republican senators [14]. Further investigation of the potential consequences of these power dynamics in the fact-checking process is a fruitful avenue of future research.

### 8.3 Implications for future research on fact-checking

Taking a multi-stakeholder perspective on the fact-checking process helped us learn the needs of stakeholder groups as well as uncover the challenges that go beyond the technical aspects of fact-checking. We use our findings to discuss and propose various directions that the future

<sup>7</sup><https://meedan.com/digital-health-lab>

<sup>8</sup>[https://www.facebook.com/journalismproject/facebook-partners-with-meedan-digital-health-lab-to-help-fact-checkers-debunk-health-misinformation?locale=pa\\_IN](https://www.facebook.com/journalismproject/facebook-partners-with-meedan-digital-health-lab-to-help-fact-checkers-debunk-health-misinformation?locale=pa_IN)

research on fact-checking can take. In Section 8.3.1, we start by discussing the needs of various stakeholder groups and propose solutions for the same. In Section 8.3.2, we discuss the values that the stakeholder groups desire in the tools and systems built for them. Next, in Section 8.3.3, we discuss how focusing on technical solutions alone is not enough and how the existing automated approaches fail to work in real life since they ignore the social aspect of fact-checking. We also reflect on the social and civic challenges faced by fact-checking organizations by discussing the role of information infrastructure on fact-checking. Finally, in Section 8.3.4, we end by stressing how the current research on misinformation has not focused on the Global South countries and how there is a dearth of fact-checking tools built for regional languages of the Global South.

*8.3.1 Technical needs of fact-checking.* Our study reveals that monitoring social media platforms is the most challenging aspect in the job of a fact-checker. A combination of over reliance on third party tools to discover potentially dubious content, limiting platform affordances, and a manual way of going through each search result to determine if its potentially misleading makes the process extremely tedious. Access to better search filters on social media platforms, and a community-based approach to reporting misinformation where users of social media platforms are able to report problematic content are some useful ways to assist fact-checkers in finding misleading claims.

Fact-checkers also agreed that its impossible for them to have access to all corners of the web. For example, they do not have access to content on private messaging platforms like WhatsApp that have become a popular haven for groups interested in sharing misinformation [33]. Given fact-checkers' skepticism towards AI and automation, user reporting via tip lines appears to be a feasible solution to access content on such platforms. Research on how to motivate people to report problematic content is a fruitful avenue of future research. We also found that fact-checkers end up viewing long videos to extract misleading claims and reading through lengthy comments section to get clues that would help them investigate the claims. A system that could utilize comments to highlight all the misleading claims present in the video, leaving the final decision of selecting what claims to verify to fact-checkers could be useful for the fact-checking community. Furthermore, a tool that highlights credibility indicators (such as comments containing useful information for investigating the claims) would significantly reduce the manual effort put in by the fact-checkers.

Our interviews revealed that when it comes to assigning a verdict to a claim there is no single standard labelling system shared across fact-checking organizations. Collective agreement on the labels could allow for greater information sharing and interoperability. Future efforts can focus on developing structured ontologies for representing credibility assessment metrics that would lead to development of common labels and benchmarks for assigning veracity labels to the claims.

Our work also sheds light on the needs of stakeholder groups other than fact-checkers. There are lack of editorial and process management tools that could be used by news desk and copy editors. Social media managers need effective strategies to increase users' engagement with the fact-checks. In order for fact-checks to really have an impact, they must be "seen and attended to by audiences" [36]. To accomplish this, it is essential to understand who shares fact-checks on social media platforms and what modality or visual storytelling technique is more suited for which platform. Tools are also needed to convert fact-checks to multiple languages to increase their reach. Platforms can also help in making the fact-checks more visible and accessible. Google's efforts (e.g. claim review) to prioritize the ranks of fact-checks in searches is one such effort. Technology critics have called structured journalism, where fact-checks are produced in a machine readable form, as the future of fact-checking [29]. Recent work has tried to automate extraction of structured information—claim, claimant, and verdict from the fact-check, to allow search engines to display it in the search results [103]. More such efforts towards structured journalism are needed to integrate fact-checks with online content.

**8.3.2 Values desired in fact-checking tools and systems.** Our study participants expressed skepticism in automation and AI technology because of its black-box nature. At the same time, they also showed willingness to adopt automated solutions for low stake tasks. Fact-checkers do not want systems that decide the veracity of information, rather they want tools that could help them with their day-to-day tasks such as monitoring online platforms, checking whether a video is digitally altered, transcribing videos in regional languages, etc. Algorithm explainability combined with tools that have humans in the loop emerged as key values that fact-checkers desire in the systems built for them. Familiarity with the inner workings of algorithms along with tools that use both human and machine capabilities for problem solving can help increase fact-checkers' trust in the automated systems. Recent times have witnessed burgeoning interest in the field of human-centered XAI where researchers draw from formal HCI theories to design explanations on how machines reach a particular decision [25, 28, 40]. Scholars are also studying how to design human and machine configurations to operationalize human-in-the-loop systems [90, 178]. Understanding specific needs for explainability with respect to fact-checkers, operationalizing those needs at the conceptual and methodological levels in tools developed for fact-checking, designing systems that could take fact-checker's feedback and use that to modify the algorithm used by the system are few useful directions for future research.

**8.3.3 Going beyond the technical: need for socio-technical solutions.** Would technology mediated solutions alone lead to improvement in the quality of fact-checking? While automating the entire fact-checking process and developing new tools to increase efficiency and scalability seems promising, it is not a panacea to all the problems faced by the stakeholder groups. A holistic change could only be achieved via systematic changes in the civic, political, and informational contexts. Through our study we found how accessing data from the civic and government bodies is a difficult task, especially if it portrays the government in a less-than-favorable light. While fact-checking is now a global endeavor, in some countries information is either not publicly available or essential information (e.g. census data, health surveys, etc.) is outdated because of lack of periodic collection. Our interviews revealed how several claims are left unchecked because of lack of sources. The availability of high quality up-to date information is essential for fact-checking—manual or automatic. Additionally, good quality data is precursor to having good machine learning models [151] that would be needed to automate the investigative step in the fact-checking pipeline where publicly available authoritative statistics are used to determine the veracity of claims. Thus, as part of *long-term advocacy centric fact-checking*, advocates within the fact-checking organizations have been actively pushing for policy changes to improve the availability and quality of data and statistics. For example, fact-checking organization Full Fact gave oral evidence to the House of Commons Public Administration and Constitutional Affairs Committee on issues of coherent and accessible health statistics in the United Kingdom [132]. There is a growing interest in the HCI community to engage with policymakers as a way to inform policy that could benefit society [114, 156]. Future work in fact-checking could focus on understanding the opportunities and difficulties that advocates face while engaging with civic organizations and determining strategies that advocates can adopt to shape policies surrounding statistics and public data in their respective countries.

**8.3.4 Fact-checking in the Global South.** Recent work in the CSCW community stressed on the fact that academic research, till date, has primarily focused on misinformation in Western countries, while not addressing the phenomenon in the Global South [93]. Our study reiterates the lack of knowledge and context surrounding the misinformation landscape in the Global South. Local regional languages are under resourced both by the online platforms and search engines making it difficult for fact-checkers to gain access to local context online and for regional speakers to gain access to reliable information [64]. The current design-based approaches to fact-checking do



not take into account the lack of fact-checking resources in regional languages and thus, fail to account for the unevenness of the viability of fact-checking across the globe. Advocators recognize how access to local community-specific knowledge and culture is essential to understand the characteristics of misinformation and why it spreads in a particular region [65]. Thus, they are calling attention to the need to improve the information infrastructure and research in the Global South. Our work also acts as call to action for researchers to study the misinformation landscape in the Global South region.

## 9 CONCLUSIONS AND LIMITATIONS

Our work sheds light on how fact-checking is practised in the real world by presenting the infrastructures—both human and technological—supporting the fact-checking work. We interviewed 26 participants belonging to six primary stakeholder groups involved in the fact-checking process namely editors, external fact-checkers, in-house fact-checkers, investigators and researchers, social media managers, and advocators. By studying the various tasks performed by these stakeholder groups, we identified the role of tools, technology and policy in their work. Finally, we also identified key challenges faced by the stakeholder groups along with opportunities of advancing current tools, policies and technology for fact-checkers.

Our work is not without limitations. The majority of the organizations that we interviewed are either IFCN signatories or work closely with IFCN signatories. Thus, the fact-checking model and tools used by the stakeholders presented by our study might not be applicable to every fact-checking organization. The fact-checking process could also be subjected to various degrees of local and regional variabilities that our work does not capture. We also acknowledge that not all stakeholder roles are present in every fact-checking organization that we interviewed. For example, not every organization is doing long term investigative or advocacy work. We leave the examination of the factors influencing the fact-checking work in different regions and the variability in fact-checking work across regions to future work.

## 10 ACKNOWLEDGEMENTS

This work was supported by National Science Foundation grant #2128642 and Fact-Checking Innovation Initiative awards. We would also like to thank all the participants without whom the study would not have been possible.

## REFERENCES

- [1] 2021. 3 Challenges of Integrating Heterogeneous Data Sources - DZone Integration. <https://dzone.com/articles/3-challenges-of-integrating-heterogeneous-data-sou>. (Accessed on 08/03/2021).
- [2] 2021. Challenges of Integrating Heterogeneous Data Sources - DATAVERSITY. <https://www.dataversity.net/challenges-of-integrating-heterogeneous-data-sources/>. (Accessed on 08/03/2021).
- [3] 2021. A Common Data Model for Europe? - Why? Which? How? - workshop report. [https://www.ema.europa.eu/en/documents/report/common-data-model-europe-why-which-how-workshop-report\\_en.pdf](https://www.ema.europa.eu/en/documents/report/common-data-model-europe-why-which-how-workshop-report_en.pdf). (Accessed on 08/03/2021).
- [4] 2021. A common data model in Europe? – Why? Which? How? | European Medicines Agency. <https://www.ema.europa.eu/en/events/common-data-model-europe-why-which-how>. (Accessed on 08/03/2021).
- [5] 2021. DER SPIEGEL | Online-Nachrichten. <https://www.spiegel.de/consent-a-?targetUrl=https%3A%2F%2Fwww.spiegel.de%2Finternational%2F&ref=https%3A%2F%2Fwww.google.com%2F>. (Accessed on 09/14/2021).
- [6] 2021. dpa: en. <https://www.dpa.com/en/>. (Accessed on 08/17/2021).
- [7] 2021. Fact Check. <https://www.indiatoday.in/fact-check>. (Accessed on 08/17/2021).
- [8] 2021. Portada · Maldita.es - Periodismo para que no te la cuelen. <https://maldita.es/>. (Accessed on 08/17/2021).
- [9] 2021. Presentación de PowerPoint. [https://maldita.es/uploads/public/docs/barometro\\_desinformacion\\_parte\\_1.pdf](https://maldita.es/uploads/public/docs/barometro_desinformacion_parte_1.pdf). (Accessed on 08/25/2021).
- [10] 2021. Search Results - LexisNexis. <https://www.lexisnexis.com/en-us/search.page>. (Accessed on 04/15/2021).
- [11] 2021. Uganda in Crisis – ANCIR's iLAB. <https://investigate.africa/reports/uganda-in-crisis/>. (Accessed on 08/25/2021).

- [12] 2022. Bringing fact check information to Google Images. <https://blog.google/products/search/bringing-fact-check-information-google-images/>. (Accessed on 01/13/2022).
- [13] 2022. Debunk Bot (@DebunkBotAfrica) / Twitter. <https://twitter.com/debunkbotafrika>. (Accessed on 01/12/2022).
- [14] 2022. Facebook Takes Down Fact-Check Of Live Action, Lila Rose Anti-Abortion Videos. <https://www.buzzfeednews.com/article/claudiakoerner/facebook-fact-check-abortion-video-doctors-medical>. (Accessed on 01/05/2022).
- [15] 2022. Gephi - The Open Graph Viz Platform. <https://gephi.org/>. (Accessed on 01/08/2022).
- [16] 2022. Google Fact Check Feature: What It Means for Your Online Efforts - Act-On. <https://act-on.com/blog/google-fact-check-feature-what-it-means-for-your-online-efforts/>. (Accessed on 01/05/2022).
- [17] 2022. HCI and the U.S. Presidential Election: A Few Thoughts on a Research Agenda | by Brent Hecht | Medium. <https://brenthecht.medium.com/hci-and-the-u-s-presidential-election-a-few-thoughts-on-a-research-agenda-7c1a0a04986>. (Accessed on 01/05/2022).
- [18] 2022. Industry-leading vector graphics software | Adobe Illustrator. <https://www.adobe.com/products/illustrator.html>. (Accessed on 01/07/2022).
- [19] 2022. Kapwing: The Collaborative Online Video Editor. <https://www.kapwing.com/>. (Accessed on 01/07/2022).
- [20] 2022. MOOC.org | Massive Open Online Courses | An edX Site. <https://www.mooc.org/>. (Accessed on 01/08/2022).
- [21] 2022. Official Adobe Photoshop | Photo and design software. <https://www.adobe.com/products/photoshop.html>. (Accessed on 01/07/2022).
- [22] 2022. See fact checks in YouTube search results - YouTube Help. <https://support.google.com/youtube/answer/9229632?hl=en>. (Accessed on 01/05/2022).
- [23] accessed in 2021. The African Network of Centers for Investigative Reporting's Investigative Lab. <https://investigate.africa/>
- [24] Ashraf Abdul, Jo Vermeulen, Danding Wang, Brian Y Lim, and Mohan Kankanhalli. 2018. Trends and trajectories for explainable, accountable and intelligible systems: An hci research agenda. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–18.
- [25] Ashraf Abdul, Jo Vermeulen, Danding Wang, Brian Y. Lim, and Mohan Kankanhalli. 2018. *Trends and Trajectories for Explainable, Accountable and Intelligible Systems: An HCI Research Agenda*. Association for Computing Machinery, New York, NY, USA, 1–18. <https://doi.org/10.1145/3173574.3174156>
- [26] Joanna Abraham and Madhu C Reddy. 2013. Re-coordinating activities: an investigation of articulation work in patient transfers. In *Proceedings of the 2013 conference on Computer supported cooperative work*. 67–78.
- [27] Mark S Ackerman. 2000. The intellectual challenge of CSCW: The gap between social requirements and technical feasibility. *Human-Computer Interaction* 15, 2-3 (2000), 179–203.
- [28] Amina Adadi and Mohammed Berrada. 2018. Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). *IEEE Access* 6 (2018), 52138–52160. <https://doi.org/10.1109/ACCESS.2018.2870052>
- [29] Bill Adair. 2021. The future of fact-checking is all about structured data. <https://www.niemanlab.org/2020/12/the-future-of-fact-checking-is-all-about-structured-data/>
- [30] Kiran Adnan, Rehan Akbar, and Khor Siak Wang. 2019. Information Extraction from Multifaceted Unstructured Big Data. *International Journal of Recent Technology and Engineering (IJRTE)* 8 (2019), 1398–1404.
- [31] AFP. 2021. Fact Check |. <https://factcheck.afp.com/>. (Accessed on 04/15/2021).
- [32] Alexander Agadjanian, Nikita Bakhru, Victoria Chi, Devyn Greenberg, Byrne Hollander, Alexander Hurt, Joseph Kind, Ray Lu, Annie Ma, Brendan Nyhan, et al. 2019. Counting the Pinocchios: The effect of summary fact-checking data on perceived accuracy and favorability of politicians. *Research & Politics* 6, 3 (2019), 2053168019870351.
- [33] Jussara Almeida. 2019. Misinformation dissemination on the web. In *Companion Proceedings of the 2019 World Wide Web Conference*. 740–740.
- [34] Majed Alrubaian, Muhammad Al-Qurishi, Mohammad Mehedi Hassan, and Atif Alamri. 2016. A credibility analysis system for assessing information on twitter. *IEEE Transactions on Dependable and Secure Computing* 15, 4 (2016), 661–674.
- [35] Michelle A Amazeen. 2013. A Critical Assessment of Fact-checking in 2012.
- [36] Michelle A Amazeen, Chris J Vargo, and Toby Hopp. 2019. Reinforcing attitudes in a gatwatching news era: Individual-level antecedents to sharing fact-checks on social media. *Communication Monographs* 86, 1 (2019), 112–132.
- [37] Internet Archive. 2021. Internet Archive: Wayback Machine. <https://archive.org/web/>. (Accessed on 04/15/2021).
- [38] Archive.today. 2021. archive.today - Wikipedia. <https://en.wikipedia.org/wiki/Archive.today>. (Accessed on 04/15/2021).
- [39] Andrew Arshnt and Daniel Etcovitch. 2018. The human cost of online content moderation. *Harvard Law Review Online, Harvard University, Cambridge, MA, USA*. Retrieved from <https://jolt.law.harvard.edu/digest/the-human-cost-of-online-content-moderation> (2018).
- [40] Vijay Arya, Rachel KE Bellamy, Pin-Yu Chen, Amit Dhurandhar, Michael Hind, Samuel C Hoffman, Stephanie Houde, Q Vera Liao, Ronny Luss, Aleksandra Mojsilović, et al. 2019. One explanation does not fit all: A toolkit and taxonomy

- of ai explainability techniques. *arXiv preprint arXiv:1909.03012* (2019).
- [41] Maryam Ashoori and Justin D Weisz. 2019. In AI we trust? Factors that influence trustworthiness of AI-infused decision-making processes. *arXiv preprint arXiv:1912.02675* (2019).
- [42] Ballotpedia. accessed in March 2021. The methodologies of fact-checking. [https://ballotpedia.org/The\\_methodologies\\_of\\_fact-checking](https://ballotpedia.org/The_methodologies_of_fact-checking)
- [43] Sérgio Barbosa and Stefania Milan. 2019. Do Not Harm in Private Chat Apps: Ethical Issues for Research on and with WhatsApp. *Westminster Papers in Communication and Culture* 14 (08 2019), 49–65. <https://doi.org/10.16997/wpcc.313>
- [44] John Carlo Bertot and Heeyoon Choi. 2013. Big Data and E-Government: Issues, Policies, and Recommendations. In *Proceedings of the 14th Annual International Conference on Digital Government Research* (Quebec, Canada) (*dg.o '13*). Association for Computing Machinery, New York, NY, USA, 1–10. <https://doi.org/10.1145/2479724.2479730>
- [45] Alessandro Bondielli and Francesco Marcelloni. 2019. A survey on fake news and rumour detection techniques. *Information Sciences* 497 (2019), 38–55.
- [46] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.
- [47] BuzzSumo. 2021. BuzzSumo.com. <https://buzzsumo.com/>. (Accessed on 04/15/2021).
- [48] Noel Cass, Tim Schwanen, and Elizabeth Shove. 2018. Infrastructures, intersections and societal transformations. *Technological Forecasting and Social Change* 137 (2018), 160–167. <https://doi.org/10.1016/j.techfore.2018.07.039>
- [49] Sylvie Cazalens, Philippe Lamarre, Julien Leblay, Ioana Manolescu, and Xavier Tannier. 2018. A content management perspective on fact-checking. In *Companion Proceedings of the The Web Conference 2018*. 565–574.
- [50] Assunta Cerone, Elham Naghizade, Falk Scholer, Devi Mallal, Russell Skelton, and Damiano Spina. 2020. Watch'n'Check: Towards a social media monitoring tool to assist fact-checking experts. In *2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA)*. IEEE, 607–613.
- [51] Bidisha Chaudhuri. 2019. Paradoxes of Intermediation in Aadhaar: Human making of a digital infrastructure. *South Asia: Journal of South Asian Studies* 42, 3 (2019), 572–587.
- [52] Africa Check. 2021. Africa Check | Sorting fact from fiction. <https://africacheck.org/>. (Accessed on 04/15/2021).
- [53] Qiang Chen, Yangyi Zhang, Richard Evans, and Chen Min. 2021. Why Do Citizens Share COVID-19 Fact-Checks Posted by Chinese Government Social Media Accounts? The Elaboration Likelihood Model. *International Journal of Environmental Research and Public Health* 18, 19 (2021), 10058.
- [54] David Cheruiyot and Raul Ferrer-Conill. 2018. “Fact-Checking Africa”. *Digital Journalism* 6, 8 (2018), 964–975. <https://doi.org/10.1080/21670811.2018.1493940> arXiv:<https://doi.org/10.1080/21670811.2018.1493940>
- [55] J. Conditt. 2017. Google partners with fact-checking network to fight fake news. <https://www.engadget.com/2017-10-26-google-fake-news-international-fact-checking-network.html>
- [56] TECH & CHECK COOPERATIVE. 2021. Tech & Check — Duke Reporters’ Lab. <https://reporterslab.org/tech-and-check/>. (Accessed on 04/15/2021).
- [57] CrowdTanle. 2021. CrowdTangle | Content Discovery and Social Monitoring Made Easy. <https://www.crowdtangle.com/>. (Accessed on 04/15/2021).
- [58] Marina Danilevsky, Kun Qian, Ranit Aharonov, Yannis Katsis, Ban Kawas, and Prithviraj Sen. 2020. A Survey of the State of Explainable AI for Natural Language Processing. In *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing*. Association for Computational Linguistics, Suzhou, China, 447–459. <https://aclanthology.org/2020.aacl-main.46>
- [59] Kevin C Desouza and Kendra L Smith. 2014. Big data for social innovation. *Stanford Social Innovation Review* 12, 3 (2014), 38–43.
- [60] Nicholas Dias and Amy Sippitt. 2020. Researching fact checking: Present limitations and future opportunities. *The Political Quarterly* 91, 3 (2020), 605–613.
- [61] Colin Dickey. 2019. The Rise and Fall of Facts. [https://www.cjr.org/special\\_report/rise-and-fall-of-fact-checking.php](https://www.cjr.org/special_report/rise-and-fall-of-fact-checking.php)
- [62] Bryan Dosono and Bryan Semaan. 2019. Moderation practices as emotional labor in sustaining online communities: The case of AAPI identity work on Reddit. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–13.
- [63] DPA. 2021. Fact Check. <https://dps-factify.com>.
- [64] First Draft. 2020. Combating misinformation in under-resourced languages: lessons from around the world. <https://firstdraftnews.org/latest/combating-misinformation-in-under-resourced-languages-lessons-from-around-the-world/>
- [65] First Draft. 2020. The importance of local context in taking on misinformation: Lessons from AfricaCheck. <https://firstdraftnews.org/latest/the-importance-of-local-context-to-taking-on-misinformation-lessons-from-africacheck/>

- [66] Catherine Dwyer. 2007. Task Technology Fit, the social technical gap and social networking sites. *AMCIS 2007 Proceedings* (2007), 374.
- [67] Michaelanne Dye, David Nemer, Josiah Mangiameli, Amy S Bruckman, and Neha Kumar. 2018. El Paquete Semanal: The Week's Internet in Havana. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–12.
- [68] Ullrich KH Ecker, Ziggy O'Reilly, Jesse S Reid, and Ee Pin Chang. 2020. The effectiveness of short-format refutational fact-checks. *British Journal of Psychology* 111, 1 (2020), 36–54.
- [69] Jane Elizabeth. 2014. Who are you calling a fact checker? <https://www.americanpressinstitute.org/fact-checking-project/fact-checker-definition/>
- [70] Ilker Etikan, Sulaiman Abubakar Musa, and Rukayya Sunusi Alkassim. 2016. Comparison of convenience sampling and purposive sampling. *American journal of theoretical and applied statistics* 5, 1 (2016), 1–4.
- [71] Full Fact. 2021. Full Fact. <https://fullfact.org/>. (Accessed on 04/15/2021).
- [72] Kaethe Post Ferguson. 2005. Impact of technology on rural Appalachian health care providers: Assessment of technological infrastructure, behaviors, and attitudes. (2005).
- [73] Jennifer Ferreira, Helen Sharp, and Hugh Robinson. 2011. User experience design and agile development: managing cooperation through articulation work. *Software: Practice and Experience* 41, 9 (2011), 963–974.
- [74] Robert S Fish, Robert E Kraut, and Mary DP Leland. 1988. Quilt: A collaborative tool for cooperative writing. In *Proceedings of the ACM SIGOIS and IEEECS TC-OA 1988 conference on Office information systems*. 30–37.
- [75] Kim Fridkin, Patrick J Kenney, and Amanda Wintersieck. 2015. Liar, liar, pants on fire: How fact-checking influences citizens' reactions to negative advertising. *Political Communication* 32, 1 (2015), 127–151.
- [76] A JOINT BRIEFING FROM. 2020. Communicating fact checks online. (2020).
- [77] FullFact. 2021. GitHub - FullFact/claim-review-schema-wordpress-plugin: An open source project to create a Wordpress plugin for claim review schema. <https://github.com/FullFact/claim-review-schema-wordpress-plugin>. (Accessed on 04/15/2021).
- [78] Huiji Gao, Xufei Wang, Geoffrey Barbier, and Huan Liu. 2011. Promoting coordination for disaster relief—from crowdsourcing to coordination. In *International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction*. Springer, 197–204.
- [79] Amira Ghenai and Yelena Mejova. 2017. Catching Zika fever: Application of crowdsourcing and machine learning for tracking health misinformation on Twitter. *arXiv preprint arXiv:1707.03778* (2017).
- [80] Fabien Girardin. 2007. *Towards Reducing the Social-Technical Gap in Location-Aware Computing*. Ph.D. Dissertation. Citeseer.
- [81] Leo A Goodman. 1961. Snowball sampling. *The annals of mathematical statistics* (1961), 148–170.
- [82] D Graves. 2018. Understanding the promise and limits of automated fact-checking. (2018).
- [83] Lucas Graves. 2013. *Deciding what's true: Fact-checking journalism and the new ecology of news*. Ph.D. Dissertation. Columbia University.
- [84] Lucas Graves. 2017. Anatomy of a fact check: Objective practice and the contested epistemology of fact checking. *Communication, Culture & Critique* 10, 3 (2017), 518–537.
- [85] Lucas Graves and Michelle A Amazeen. 2019. Fact-checking as idea and practice in journalism. In *Oxford Research Encyclopedia of Communication*.
- [86] Lucas Graves and Charles W Anderson. 2020. Discipline and promote: Building infrastructure and managing algorithms in a “structured journalism” project by professional fact-checking groups. *New Media & Society* 22, 2 (2020), 342–360.
- [87] Lucas Graves and Federica Cherubini. 2016. The rise of fact-checking sites in Europe. (2016).
- [88] Lucas Graves and Tom Glaisyer. 2012. The fact-checking universe in Spring 2012. *New America* (2012).
- [89] Rebecca E Grinter. 1995. Using a configuration management tool to coordinate software development. In *Proceedings of conference on Organizational computing systems*. 168–177.
- [90] Tor Grønsvund and Margunn Aanestad. 2020. Augmenting the algorithm: Emerging human-in-the-loop work configurations. *The Journal of Strategic Information Systems* 29, 2 (2020), 101614. <https://doi.org/10.1016/j.jsis.2020.101614> Strategic Perspectives on Digital Work and Organizational Transformation.
- [91] F Guerra, D Linz, R Garcia, B Kommatata, J Kosiuk, J Chun, S Boveda, and D Duncker. 2021. The use of instant messaging in clinical data sharing: the EHRA SMS survey. *EP Europace* 23, Supplement\_3 (2021), eua116–515.
- [92] Peter H. Carstensen. 1995. Modeling coordination work: Lessons learned from analyzing a cooperative work setting. In *Symbiosis of Human and Artifact*, Yuichiro Anzai, Katsuhiko Ogawa, and Hirohiko Mori (Eds.). Advances in Human Factors/Ergonomics, Vol. 20. Elsevier, 327–332. [https://doi.org/10.1016/S0921-2647\(06\)80054-7](https://doi.org/10.1016/S0921-2647(06)80054-7)
- [93] Md Mahfuzul Haque, Mohammad Yousuf, Ahmed Shatil Alam, Pratyasha Saha, Syed Ishtiaque Ahmed, and Naeemul Hassan. 2020. Combating Misinformation in Bangladesh: Roles and Responsibilities as Perceived by Journalists, Fact-checkers, and Users. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW2 (2020), 1–32.

- [94] Naemul Hassan, Fatma Arslan, Chengkai Li, and Mark Tremayne. 2017. Toward automated fact-checking: Detecting check-worthy factual claims by claimbuster. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 1803–1812.
- [95] Naemul Hassan, Mohammad Yousuf, Md Mahfuzul Haque, Javier A. Suarez Rivas, and Md Khadimul Islam. 2019. Examining the roles of automation, crowds and professionals towards sustainable fact-checking. In *Companion Proceedings of The 2019 World Wide Web Conference*. 1001–1006.
- [96] Fazhi He and Soonhung Han. 2006. A method and tool for human–human interaction and instant collaboration in CSCW-based CAD. *Computers in Industry* 57, 8-9 (2006), 740–751.
- [97] Richard Heeks. 2003. Most eGovernment-for-development projects fail: how can risks be reduced? (2003).
- [98] Kyle Hunt, Puneet Agarwal, and Jun Zhuang. 2020. Monitoring misinformation on Twitter during crisis events: a machine learning approach. *Risk analysis* (2020).
- [99] Influencer. 2021. Find Everything About YouTube on NoxInfluencer. <https://www.noxinfluencer.com/>. (Accessed on 04/15/2021).
- [100] InVID. 2021. InVID Verification Plugin - InVID project. <https://www.invid-project.eu/tools-and-services/invid-verification-plugin/>. (Accessed on 04/15/2021).
- [101] Margaret Jack, Jay Chen, and Steven J Jackson. 2017. Infrastructure as creative action: Online buying, selling, and delivery in Phnom Penh. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 6511–6522.
- [102] Al Jazeera. 2021. Breaking News, World News and Video from Al Jazeera | Today's latest from Al Jazeera. <https://www.aljazeera.com/>. (Accessed on 04/15/2021).
- [103] Shan Jiang, Simon Baumgartner, Abe Ittycheriah, and Cong Yu. 2020. Factoring fact-checks: Structured information extraction from fact-checking articles. In *Proceedings of The Web Conference 2020*. 1592–1603.
- [104] Marina Jirotko, Rob Procter, Tom Rodden, and Geoffrey C Bowker. 2006. Collaboration in e-Research. *Computer Supported Cooperative Work (CSCW)* 15, 4 (2006), 251–255.
- [105] Georgi Karadzhov, Preslav Nakov, Lluís Màrquez, Alberto Barrón-Cedeño, and Ivan Koychev. 2017. Fully automated fact checking using external sources. *arXiv preprint arXiv:1710.00341* (2017).
- [106] Helena Karasti and Jeanette Blomberg. 2018. Studying infrastructuring ethnographically. *Computer Supported Cooperative Work (CSCW)* 27, 2 (2018), 233–265.
- [107] Aaron Karp and Bryan Pardo. 2017. HaptEQ: A collaborative tool for visually impaired audio producers. In *Proceedings of the 12th International Audio Mostly Conference on Augmented and Participatory Sound and Music Experiences*. 1–4.
- [108] Yavuz Selim Kartal, Busra Guvenen, and Mucahid Kutlu. 2020. Too many claims to fact-check: Prioritizing political claims based on check-worthiness. *arXiv preprint arXiv:2004.08166* (2020).
- [109] Ashkan Kazemi, Kiran Garimella, Gautam Kishore Shahi, Devin Gaffney, and Scott A Hale. 2021. Tiplines to Combat Misinformation on Encrypted Platforms: A Case Study of the 2019 Indian Election on WhatsApp. *arXiv preprint arXiv:2106.04726* (2021).
- [110] Aphra Kerr and John D Kelleher. 2015. The recruitment of passion and community in the service of capital: Community managers in the digital games industry. *Critical studies in media communication* 32, 3 (2015), 177–192.
- [111] Jooyeon Kim, Behzad Tabibian, Alice Oh, Bernhard Schölkopf, and Manuel Gomez-Rodriguez. 2018. Leveraging the crowd to detect and reduce the spread of fake news and misinformation. In *Proceedings of the eleventh ACM international conference on web search and data mining*. 324–332.
- [112] Neema Kotonya and Francesca Toni. 2020. Explainable automated fact-checking for public health claims. *arXiv preprint arXiv:2010.09926* (2020).
- [113] Airi Lampinen, Victoria Bellotti, Coye Cheshire, and Mary Gray. 2016. CSCW and the Sharing Economy: The Future of Platforms as Sites of Work Collaboration and Trust. In *Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion*. 491–497.
- [114] Jonathan Lazar. 2015. Public Policy and HCI: Making an Impact in the Future. *Interactions* 22, 5 (Aug. 2015), 69–71. <https://doi.org/10.1145/2807916>
- [115] Data Leads. 2021. dataleads. <https://dataleads.co.in/>. (Accessed on 04/15/2021).
- [116] Julien Leblay, Ioana Manolescu, and Xavier Tannier. 2018. Computational fact-checking: Problems, state of the art, and perspectives. In *The Web Conference*.
- [117] Charlotte P Lee, Paul Dourish, and Gloria Mark. 2006. The human infrastructure of cyberinfrastructure. In *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*. 483–492.
- [118] Pantelis Linardatos, Vasilis Papastefanopoulos, and Sotiris Kotsiantis. 2021. Explainable ai: A review of machine learning interpretability methods. *Entropy* 23, 1 (2021), 18.
- [119] Nina Lundberg and Hilda Tellioglu. 1999. Understanding complex coordination processes in health care. *Scandinavian Journal of Information Systems* 11, 1 (1999), 5.
- [120] Gloria Mark, Ban Al-Ani, and Bryan Semaan. 2009. Repairing human infrastructure in war zones. *Proceedings of ISCRAM* (2009), 10–13.



- [121] Randi Markussen. 1996. Politics of intervention in design: Feminist reflections on the Scandinavian tradition. *ai & Society* 10, 2 (1996), 127–141.
- [122] Maurice Mars and Richard E Scott. 2016. Whatsapp in clinical practice: A literature. *The Promise of New Technologies in an Age of New Health Challenges* (2016), 82.
- [123] Meedan. 2021. Meedan. <https://meedan.com/>. (Accessed on 04/15/2021).
- [124] Philipe Melo, Johnatan Messias, Gustavo Resende, Kiran Garimella, Jussara Almeida, and Fabrício Benevenuto. 2019. WhatsApp Monitor: A Fact-Checking System for WhatsApp. *Proceedings of the International AAAI Conference on Web and Social Media* 13, 01 (Jul. 2019), 676–677. <https://ojs.aaai.org/index.php/ICWSM/article/view/3271>
- [125] Tim Miller. 2019. Explanation in artificial intelligence: Insights from the social sciences. *Artificial intelligence* 267 (2019), 1–38.
- [126] Jonathan T Morgan, Michael Gilbert, David W McDonald, and Mark Zachry. 2013. Project talk: Coordination work and group membership in WikiProjects. In *Proceedings of the 9th International Symposium on Open Collaboration*. 1–10.
- [127] Dulce Mota, Carlos Vaz de Carvalho, and Luis Paulo Reis. 2011. Fostering Collaborative Work between educators in higher education. In *2011 IEEE International Conference on Systems, Man, and Cybernetics*. IEEE, 1286–1291.
- [128] Bonnie A Nardi and Yrjö Engeström. 1999. A web on the wind: The structure of invisible work. *Computer supported cooperative work* 8, 1-2 (1999), 1–8.
- [129] An T Nguyen, Aditya Kharosekar, Saumyaa Krishnan, Siddhesh Krishnan, Elizabeth Tate, Byron C Wallace, and Matthew Lease. 2018. Believe it or not: Designing a human-ai partnership for mixed-initiative fact-checking. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*. 189–199.
- [130] Lilly U Nguyen. 2016. Infrastructural action in Vietnam: Inverting the techno-politics of hacking in the global South. *New Media & Society* 18, 4 (2016), 637–652.
- [131] Anne Oeldorf-Hirsch, Mike Schmierbach, Alyssa Appelman, and Michael P Boyle. 2020. The ineffectiveness of fact-checking labels on news memes and articles. *Mass Communication and Society* 23, 5 (2020), 682–704.
- [132] House of Commons. 2019. Public Administration and Constitutional Affairs Committee, Oral evidence: Governance of Statistics. <http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/public-administration-and-constitutional-affairs-committee/governance-of-official-statistics/oral/97624.html>
- [133] Andrés Páez. 2019. The pragmatic turn in explainable artificial intelligence (XAI). *Minds and Machines* 29, 3 (2019), 441–459.
- [134] Sachin R Pendse, Faisal M Lalani, Munmun De Choudhury, Amit Sharma, and Neha Kumar. 2020. " Like Shock Absorbers": Understanding the Human Infrastructures of Technology-Mediated Mental Health Support. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–14.
- [135] Pesacheck. 2021. PesaCheck. <https://pesacheck.org/>. (Accessed on 04/15/2021).
- [136] The Washington Post. 2021. About The Fact Checker - The Washington Post. <https://www.washingtonpost.com/politics/2019/01/07/about-fact-checker/>. (Accessed on 04/15/2021).
- [137] The Washington Post. 2021. Fact Checker - The Washington Post. <https://www.washingtonpost.com/news/fact-checker/>. (Accessed on 04/15/2021).
- [138] Poynter. accessed in March, 2021. The International Fact-Checking Network. <https://www.poynter.org/ifcn/>
- [139] The Quint. 2021. Latest News, Breaking News LIVE, Top News Headlines, Viral Videos News Updates - The Quint. <https://www.thequint.com/>. (Accessed on 04/15/2021).
- [140] Lauge Baungaard Rasmussen. 2007. From human-centred to human-context centred approach: looking back over ‘the hills’, what has been gained and lost? *Ai & Society* 21, 4 (2007), 471–495.
- [141] The New Republic. 2021. The New Republic. <https://newrepublic.com/>. (Accessed on 04/15/2021).
- [142] Paul Resnick, Samuel Carton, Souneil Park, Yuncheng Shen, and Nicole Zeffer. 2014. Rumorlens: A system for analyzing the impact of rumors and corrections in social media. In *Proc. Computational Journalism Conference*, Vol. 5.
- [143] Sarah T Roberts. 2014. *Behind the screen: The hidden digital labor of commercial content moderation*. Ph.D. Dissertation. University of Illinois at Urbana-Champaign.
- [144] John J Robinson, Jim Maddock, and Kate Starbird. 2015. Examining the Role of Human and Technical Infrastructure during Emergency Response.. In *ISCRAM*.
- [145] Nikol Rummel, Hans Spada, Fabian Hermann, Franz Caspar, and Katrin Schornstein. 2002. Promoting the coordination of computer-mediated interdisciplinary collaboration. (2002).
- [146] Nithya Sambasivan and Thomas Smyth. 2010. The human infrastructure of ICTD. In *Proceedings of the 4th ACM/IEEE international conference on information and communication technologies and development*. 1–9.
- [147] Steve Sawyer and Andrea Tapia. 2006. Always articulating: Theorizing on mobile and wireless technologies. *The Information Society* 22, 5 (2006), 311–323.
- [148] Nicolas Scharowski. 2020. *Transparency and Trust in AI*. Ph.D. Dissertation. Institute of Psychology.



- [149] Philipp Schmidt, Felix Biessmann, and Timm Teubner. 2020. Transparency and trust in artificial intelligence systems. *Journal of Decision Systems* 29, 4 (2020), 260–278.
- [150] Social Searcher. 2021. Social Searcher - Free Social Media Search Engine. <https://www.social-searcher.com/>. (Accessed on 04/15/2021).
- [151] Valerie Sessions and Marco Valtorta. 2006. The Effects of Data Quality on Machine Learning Algorithms. *ICIQ* 6 (2006), 485–498.
- [152] Chengcheng Shao, Giovanni Luca Ciampaglia, Alessandro Flammini, and Filippo Menczer. 2016. Hoaxy: A platform for tracking online misinformation. In *Proceedings of the 25th international conference companion on world wide web*. 745–750.
- [153] Prashant Shiralkar, Alessandro Flammini, Filippo Menczer, and Giovanni Luca Ciampaglia. 2017. Finding streams in knowledge graphs to support fact checking. In *2017 IEEE International Conference on Data Mining (ICDM)*. IEEE, 859–864.
- [154] Susan Currie Sivek and Sharon Bloyd-Peshkin. 2019. Where Do Facts Matter? The Digital Paradox in Magazines’ Fact-checking Processes. *Journalism Practice* 13, 8 (2019), 998–1002. <https://doi.org/10.1080/17512786.2019.1643767> arXiv:<https://doi.org/10.1080/17512786.2019.1643767>
- [155] Alison Smith, Varun Kumar, Jordan Boyd-Graber, Kevin Seppi, and Leah Findlater. 2018. Closing the loop: User-centered design and evaluation of a human-in-the-loop topic modeling system. In *23rd International Conference on Intelligent User Interfaces*. 293–304.
- [156] Anne Spaa, Abigail Durrant, Chris Elsdén, and John Vines. 2019. Understanding the Boundaries between Policymaking and HCI. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) (CHI ’19). Association for Computing Machinery, New York, NY, USA, 1–15. <https://doi.org/10.1145/3290605.3300314>
- [157] Spoonbill. 2021. Spoonbill. <https://spoonbill.io/>. (Accessed on 04/15/2021).
- [158] Susan Leigh Star and Karen Ruhleder. 1996. Steps toward an ecology of infrastructure: Design and access for large information spaces. *Information systems research* 7, 1 (1996), 111–134.
- [159] Susan Leigh Star and Anselm Strauss. 1999. Layers of silence, arenas of voice: The ecology of visible and invisible work. *Computer supported cooperative work (CSCW)* 8, 1 (1999), 9–30.
- [160] Miriah Steiger, Timir J Bharucha, Sukrit Venkatagiri, Martin J Riedl, and Matthew Lease. 2021. The Psychological Well-Being of Content Moderators. (2021).
- [161] Mark Stencil. 2019. Number of fact-checking outlets surges to 188 in more than 60 countries. <https://www.poynter.org/fact-checking/2019/number-of-fact-checking-outlets-surges-to-188-in-more-than-60-countries/>
- [162] Allan Stisen, Nervo Verdezoto, Henrik Blunck, Mikkel Baun Kjærgaard, and Kaj Grønbaek. 2016. Accounting for the invisible work of hospital orderlies: Designing for local and global coordination. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. 980–992.
- [163] Charlotte Tang, Yunan Chen, Bryan C Semaan, and Jahmeilah A Roberson. 2015. Restructuring human infrastructure: The impact of EHR deployment in a volunteer-dependent clinic. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. 649–661.
- [164] Denis Teyssou, Jean-Michel Leung, Evlampios Apostolidis, Konstantinos Apostolidis, Symeon Papadopoulos, Markos Zampoglou, Olga Papadopoulou, and Vasileios Mezaris. 2017. The InVID plug-in: web video verification on the browser. In *Proceedings of the first international workshop on multimedia verification*. 23–30.
- [165] James Thorne, Andreas Vlachos, Christos Christodoulopoulos, and Arpit Mittal. 2018. Fever: a large-scale dataset for fact extraction and verification. *arXiv preprint arXiv:1803.05355* (2018).
- [166] Peter Tolmie, Rob Procter, David William Randall, Mark Rouncefield, Christian Burger, Geraldine Wong Sak Hoi, Arkaitz Zubiaga, and Maria Liakata. 2017. Supporting the use of user generated content in journalistic practice. In *Proceedings of the 2017 chi conference on human factors in computing systems*. 3632–3644.
- [167] Trello. 2021. Trello. <https://trello.com/en-US>. (Accessed on 04/15/2021).
- [168] David Tuffley. 2009. Mind the gap. *International Journal of Sociotechnology and Knowledge Development (IJSKD)* 1, 1 (2009), 58–69.
- [169] Kenton T Unruh and Wanda Pratt. 2008. The Invisible Work of Being a Patient and Implications for Health Care: “[the doctor is] my business partner in the most important business in my life, staying alive.” In *Ethnographic Praxis in Industry Conference Proceedings*, Vol. 2008. Wiley Online Library, 40–50.
- [170] Marten van der Meulen and W Gudrun Reijnierse. 2020. FactCorp: A Corpus of Dutch Fact-checks and its Multiple Usages. In *Proceedings of The 12th Language Resources and Evaluation Conference*. 1286–1292.
- [171] Nervo Verdezoto, Naveen Bagalkot, Syeda Zainab Akbar, Swati Sharma, Nicola Mackintosh, Deirdre Harrington, and Paula Griffiths. 2021. The Invisible Work of Maintenance in Community Health: Challenges and Opportunities for Digital Health to Support Frontline Health Workers in Karnataka, South India. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021), 1–31.

- [172] Nicholas Vincent, Brent Hecht, and Shilad Sen. 2019. "Data Strikes": Evaluating the Effectiveness of a New Form of Collective Action Against Technology Companies. In *The World Wide Web Conference*. 1931–1943.
- [173] Nicholas Vincent, Isaac Johnson, Patrick Sheehan, and Brent Hecht. 2019. Measuring the importance of user-generated content to search engines. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 13. 505–516.
- [174] William Yang Wang. 2017. " liar, liar pants on fire": A new benchmark dataset for fake news detection. *arXiv preprint arXiv:1705.00648* (2017).
- [175] Barbara Wasson. 1998. Identifying coordination agents for collaborative telelearning. *International Journal of Artificial Intelligence in Education (IJAIED)* 9 (1998), 275–299.
- [176] Merriam Webster. 2022. Stringer Definition & Meaning - Merriam-Webster. <https://www.merriam-webster.com/dictionary/stringer>. (Accessed on 04/17/2022).
- [177] Whois.net. 2021. Whois Lookup & IP | Whois.net. <https://www.whois.net/>. (Accessed on 04/15/2021).
- [178] Doris Xin, Litian Ma, Jialin Liu, Stephen Macke, Shuchen Song, and Aditya Parameswaran. 2018. Accelerating human-in-the-loop machine learning: Challenges and opportunities. In *Proceedings of the second workshop on data management for end-to-end machine learning*. 1–4.
- [179] Dannagal G Young, Kathleen Hall Jamieson, Shannon Poulsen, and Abigail Goldring. 2018. Fact-checking effectiveness as a function of format and tone: Evaluating FactCheck. org and FlackCheck. org. *Journalism & Mass Communication Quarterly* 95, 1 (2018), 49–75.

Received January 2022; revised April 2022; accepted May 2022