

Identification Responses

4

Chapter 4 of the report: **Balancing Act: Countering Digital Disinformation While Respecting Freedom of Expression**

Broadband Commission research report
on 'Freedom of Expression and Addressing
Disinformation on the Internet'

Published in 2020 by International Telecommunication Union (ITU), Place des Nations, CH-1211 Geneva 20, Switzerland, and the United Nations Educational, Scientific and Cultural Organization, and United Nations Educational, Scientific and Cultural Organization (UNESCO), 7, Place de Fontenoy, 75352 Paris 07 SP, France

ISBN 978-92-3-100403-2



This research will be available in Open Access under the Attribution-ShareAlike 3.0 IGO (CC-BY SA 3.0 IGO) license. By using the content of this publication, the users accept to be bound by the terms of use of the UNESCO Open Access Repository

<https://en.unesco.org/publications/balanceact>

4.1 Monitoring and fact-checking responses

Authors: Denis Teyssou, Julie Posetti and Sam Gregory

This chapter maps out and systematises fact-checking responses as part of monitoring disinformation (alongside investigative responses to disinformation, as covered in chapter 4.2). Here, the emphasis is on fact-checking responses which are global, regional and national in scope in a wide range of the countries and languages, and they can be either independent operations or affiliated with news organisations. The way these efforts engage in countering disinformation is also described in this chapter.

Definitions

The discipline of verification has been described as the essence of journalism (Kovach & Rosenstiel, 2001). Verification is an editorial technique used by journalists and by independent fact-checkers to verify the accuracy of a statement, and/or documents and other artifacts, as well as the platforms and identities (human and digital) of those producing or transmitting content. But there are distinctions to be drawn between verification and fact-checking (Silverman et al., 2014):

- **Verification** *is a discipline that lies at the heart of journalism, and that is increasingly being practiced and applied by other professions.*
- **Fact checking** *is a specific application of verification - both within journalism [and by other organisations, including NGOs]. In this respect, verification is a fundamental practice that enables fact checking.*

Increasingly, fact-checking also involves a process of proactive de-bunking - i.e publishing debunks to demonstrate falsehoods, and often by setting out the systematic process involved in reaching this conclusion.

4.1.1 What and who do they target?

Fact-checking responses consist of applying verification not only to the process of journalistic work (and its outputs), but also to third-party claims, statements and datasets circulating outside the legacy media sphere, especially on social networks.

Verifying the authenticity of an actor, institution or a social media account is where fact-checking begins to feed into investigative responses to disinformation (see Chapter 4.2). Identification responses, like monitoring and fact-checking, underpin the investigations into the origins and spread of disinformation, contributing to the evidence-base upon which other types of disinformation responses depend.

Specific examples will be provided in section 4.1.4 below.

4.1.2 Who do monitoring and fact-checking responses try to help?

The usefulness of fact-checking for internet communications companies⁴⁰ enables them to identify disinformation and develop responses that reduce or remove its visibility and/or credibility. Checking also helps governments and international organisations to decide what, when and whether action needs to be taken - for instance, launching policy or practical initiatives like targeted counter-disinformation campaigns. Finally, published fact-checks provide a useful source of authoritative information for citizens.

4.1.3 What output do they publish?

This response publishes its findings - what was checked, how, and what the status is in terms of validity or falsity, indeterminate or other (e.g. opinion - which is not fact-checkable per se, although where it is justified on the basis of purported facts, these aspects are prima facie checkable concerning the extent to which such 'facts' are false or misleading). It is recognised that published fact checks tend to attract fewer user shares on social media than the viral disinformation they are debunking (Shin & Thorsen, 2017). There is also some concern that drawing attention to falsehoods can help amplify them. Nevertheless, the operating assumption is that the work of verification and debunking remains essential as a means for surfacing truth and for holding individuals, public figures, institutions and the media accountable for inaccurate claims (Sippitt, 2020; Friedman, 2020; Qui, 2020).

4.1.4 Who are the primary actors and who funds these responses?

a. Global responses

First Draft

One of the early global initiatives focused on social media content verification at the international level is the non-profit coalition First Draft, registered in the UK since June 2015. The aim of First Draft at its establishment was to provide practical and ethical guidance to the news media on identifying, verifying and publishing content that is sourced from the social web, especially in breaking news contexts.

In September 2016, the original nucleus of nine partners (BellingCat, Dig Deeper, Emergent.info, EyeWitness Media Hub, Google News Initiative, Meedan, [Reported.ly](#), Storyful, and VerificationJunkie) expanded to an international Partner Network of media organisations, academics, social network platforms and civil society organisations. At the same time, First Draft joined ProPublica's project ElectionLand, which aimed to identify and track voters' encounters with misinformation and disinformation during the 2016 U.S. presidential election. They worked collaboratively on this project with students from 13 journalism schools who were trained in social newsgathering and verification techniques. Electionland was financially supported by Google News Lab and U.S. philanthropist Craig Newmark.

⁴⁰ <https://www.disinfo.eu/resources/covid-19/platforms-responses-to-covid-19-mis-and-disinformation;>

Next, First Draft launched several collaborative election-monitoring programs in France, the United Kingdom, Germany, Brazil, and Nigeria. The resulting news media and fact-checking coalition, known as CrossCheck, monitors rumours being spread within these countries, and publishes debunks of false information in order to give voters the means to reach conclusions by themselves without being misdirected by disinformation. (For more on election-targeted responses, see section 5.3).

In 2020, First Draft was expanding operations in Argentina, Australia, Canada, Indonesia, South Africa, Spain and Uruguay, and aiming to coordinate a cross-border project to investigate misinformation tactics and trends in Europe beyond election periods.

Apart from founding partner Google News Initiative, First Draft has also obtained grants and donations from many philanthropic foundations as well as support from the Facebook Journalism Project and Twitter. After briefly joining the Shorenstein Center for Media, Politics and Public Policy at Harvard's Kennedy School in October 2017, First Draft is now operating independently again, primarily relying on funding from internet communications companies. More First Draft collaborative initiatives around elections are detailed in section 5.3.

International Fact Checking Network (IFCN)

The International Fact-Checking Network (IFCN, 2019a) was launched in September 2015 as a business unit within the non-profit journalism school Poynter Institute for Media Studies, based in St. Petersburg, Florida, U.S.. The Institute, which owns the *Tampa Bay Times*, launched IFCN to bring together fact-checkers worldwide and to promote good practices and knowledge exchange in the field.

The IFCN's mission is to monitor trends, formats and policy-making about fact-checking worldwide, to publish regular articles about fact-checking, to promote training - both in person and online - as well as ensuring basic standards through the fact-checkers' code of principles. On August 6th, 2020, IFCN had 79 verified active signatories of its code of principles, 14 verified signatories under renewal (IFCN, 2020d). The map below shows the geographic distribution of the signatories. Some of them are fact-checking both in their homelands and across international borders.

A verification process is important because it is possible that in this contested terrain that flawed, or even fake, fact-checking initiatives can exploit the label for purposes far removed from challenging falsehoods.

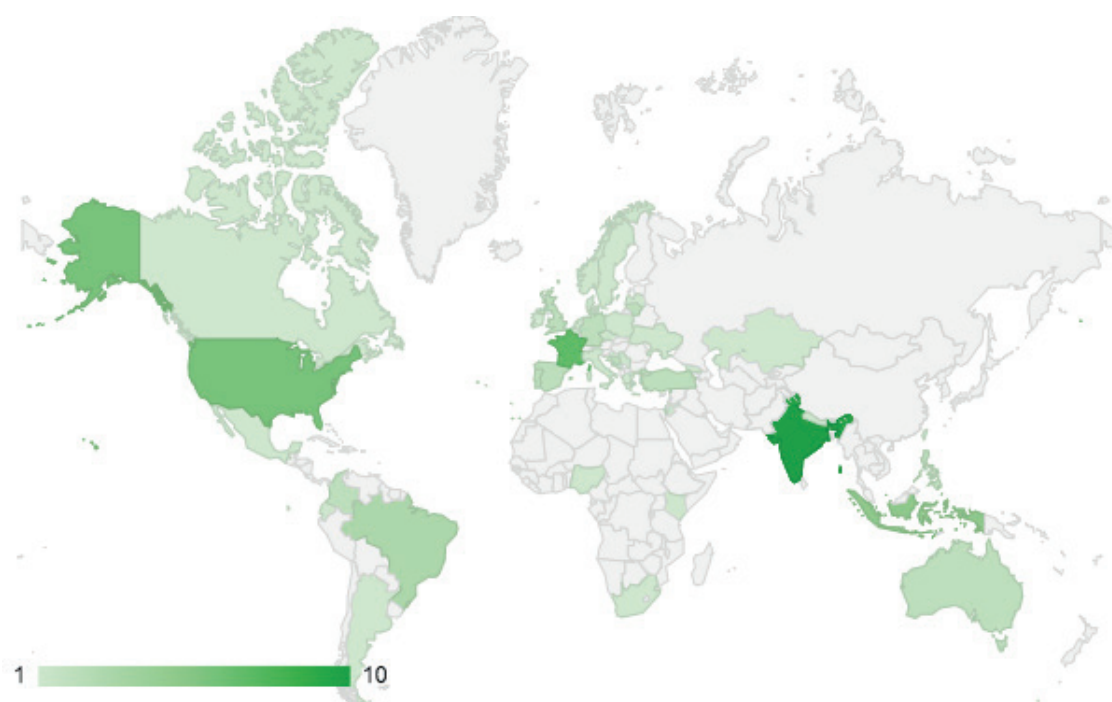


Figure 4. A geographical map of the IFCN signatories (67 verified active and 14 under renewal in early 2020)

Signatories must abide by five commitments (IFCN, 2019c):

1. A commitment to Nonpartisanship and Fairness
2. Transparency of Sources
3. Transparency of Funding and Organisation
4. Transparency of Methodology
5. Open and Honest Corrections Policy

This code of principles was launched in September 2016, one year after the birth of the IFCN. In 2017, the IFCN introduced an application and vetting process following the announcement by Facebook that being a signatory to this code is a minimum condition for being accepted as a third-party fact-checker for the company.⁴¹

Transparency, often presented in media studies literature as a new ethical tenet of journalism, plays an important role in these commitments. This intersects with the emergence of transparency among bloggers and early fact-checkers as a necessary or natural alternative to the professional journalistic ideal of objectivity (Graves, 2013). It builds on an idea espoused by philosopher David Weinberger in 2009: “transparency is the new objectivity” (Weinberger, 2009). The notion of transparency and its connection to trust in credible journalism is now widely embedded as a norm within both fact-checking operations and professional journalism. The transparency afforded by published explanations of verification and fact-checking processes can make the work more defensible against claims of bias or inaccuracy because the evidentiary base of the fact-checking exercise is laid bare.

⁴¹ See chapters 4.2 and 7.1 for further discussion

The IFCN organises a yearly international conference (Global Fact) promoting collaborative efforts between fact-checkers all over the World. The 2019 edition, Global Fact 6, was staged in Cape Town, South Africa, with more than 250 participants representing 55 countries and 146 active organisations. Global Fact 7, which was due to be held in Oslo, Norway, in June 2020, was ultimately held online due to the COVID-19 pandemic. The Network also funds annual fellowships, a Fact Forward Fund, a Fact-Checking Innovation Initiative and a crowdfunding match program. Finally IFCN advocates for a global expansion of fact-checking efforts, including through an annual International Fact-Checking Day, every April 2.

IFCN has received funding from the Arthur M. Blank Family Foundation, the Duke Reporters’ Lab, the Bill & Melinda Gates Foundation, Google, the National Endowment for Democracy, the Omidyar Network, the Open Society Foundations and the Park Foundation (IFCN, 2019d).

Duke University Reporter’s Lab database

The Reporters’ Lab is a centre for journalism research in the Sanford School of Public Policy at Duke University in the U.S.. One of its main projects has been to create a [worldwide database](https://reporterslab.org/fact-checking/)⁴² of the main fact-checking operations, active or inactive, and therefore to document the rise of the fact-checking sector, country by country.

Aside from a geographical mashup displaying all fact-checking organisations, the database allows the user to browse the content by continents and countries and it is regularly updated. Criteria to add new fact-checking sites include non-partisanship, an emphasis on reviewing fulfilment of political promises (e.g. party manifestos during elections), transparency about sources and methods, transparency about funding and affiliations, and a primary mission being news and information. As of April 2020, the Reporters’ Lab database included 237 active sites and 91 inactive worldwide in 78 countries.

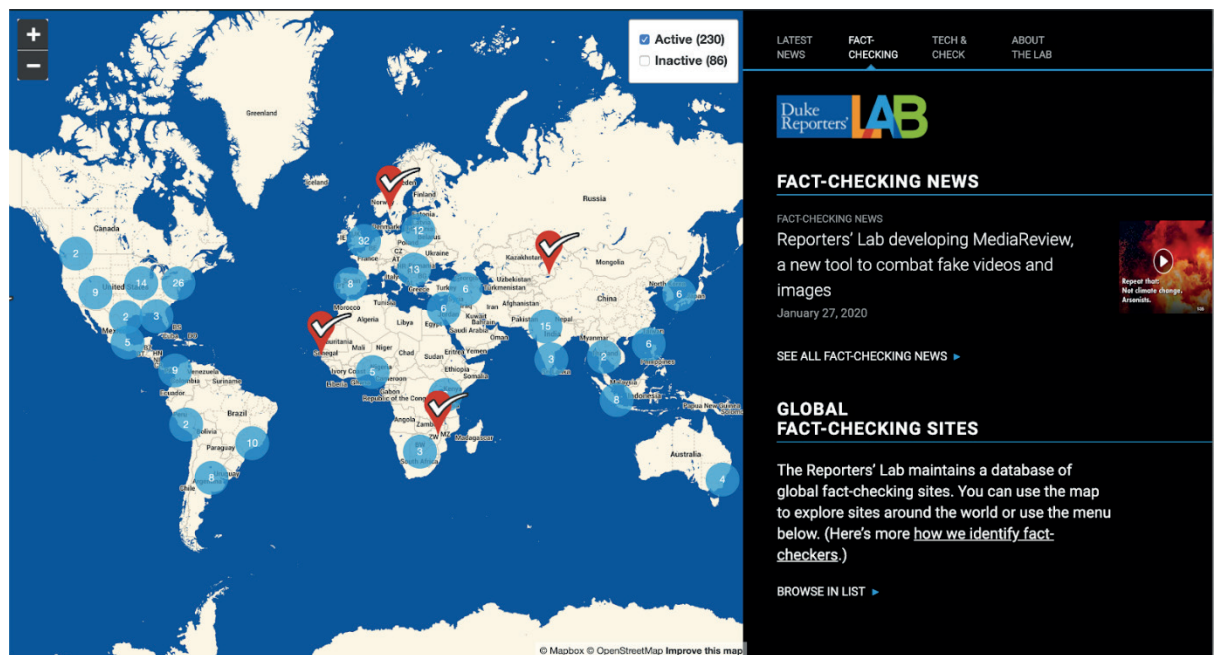


Figure 5. A view of the Duke University Reporters’ Lab fact-checking database

⁴² <https://reporterslab.org/fact-checking/>

Facebook Third-Party Fact Checking network

Internet communications companies typically employ internal or external fact-checking processes, which inform their curatorial responses, e.g. removal, demotion, or hiding of posts. These are described in more detail in Chapter 6.1.

Among the companies' systems, Facebook merits attention as the only large-scale international "third party verification" programme among the internet communications companies, which was launched shortly after the 2016 U.S. presidential election (Zuckerberg, 2016a). Announcing the project on his own Facebook page, CEO Mark Zuckerberg stated that Facebook was taking "misinformation seriously," and acknowledged that there were many respected fact-checking organisations that his company was planning to learn from. Previously, he had stated that more than 99% of what people see on Facebook is authentic (Zuckerberg, 2016b). His announcement of the third party fact-checking initiative was widely interpreted as an attempt to counter criticism of the company's lack of intervention to stem the spread of disinformation during the 2016 U.S. presidential election campaign.

One month after the U.S. election in 2016, Facebook announced the launch of a programme to work with third-party fact checking organisations who were signatories of the Poynter Institute's International Fact Checking Network (IFCN) Code of Principles (Mosseri, 2016). The Facebook third party fact-checking program sub-contracts organisations to review and rate the accuracy of content, including stories and non-political advertising (See discussion below about the limitations applied to fact-checking political content under this program).

Once a story is rated as false by these fact-checking partners, Facebook shows it lower in the 'Newsfeed' unless it is revised by Facebook in light of their policies, processes and/or payments associated with the contracts under which the fact-checking organisations operate (Pasternack, 2020). (Generally opinion content, and generally certain categories of political advertising and political speech from politicians, political parties and affiliates are excluded). On Instagram, Facebook makes content flagged under this program harder to find by filtering it from Explore and hashtag pages, and downranking it in the feed. In addition, content across Facebook and Instagram that has been rated false or partly false is prominently labelled⁴³ so people can better decide for themselves what to read, trust, and share. These labels are shown on top of false and partly false photos and videos, including on top of 'Stories' content on Instagram, and link out to the assessment from the fact-checker.⁴⁴

Prior to December 2017, if a fact-checking organisation identified a story as false (or 'fake' according to Facebook's protocol), they reported it to Facebook and it was flagged as disputed, with a link to the corresponding article (on fact-checker's website) explaining why.

According to Facebook, this limits the visibility of such posts by 80% (Lyons, 2018a) and therefore helps contain its spread. However, this can take up to three days after the content is first published (Silverman, 2017b). Facebook says that it also uses the information from fact-checkers in order to improve its technology to identify false content faster. Further assessment of labelling can be found in chapter 7.3.

⁴³ <https://about.fb.com/wp-content/uploads/2020/06/Elections-Fact-Sheet.pdf>

⁴⁴ <https://about.fb.com/news/2019/10/update-on-election-integrity-efforts/>

Facebook 3rd Party Fact-checking network by continent and operations

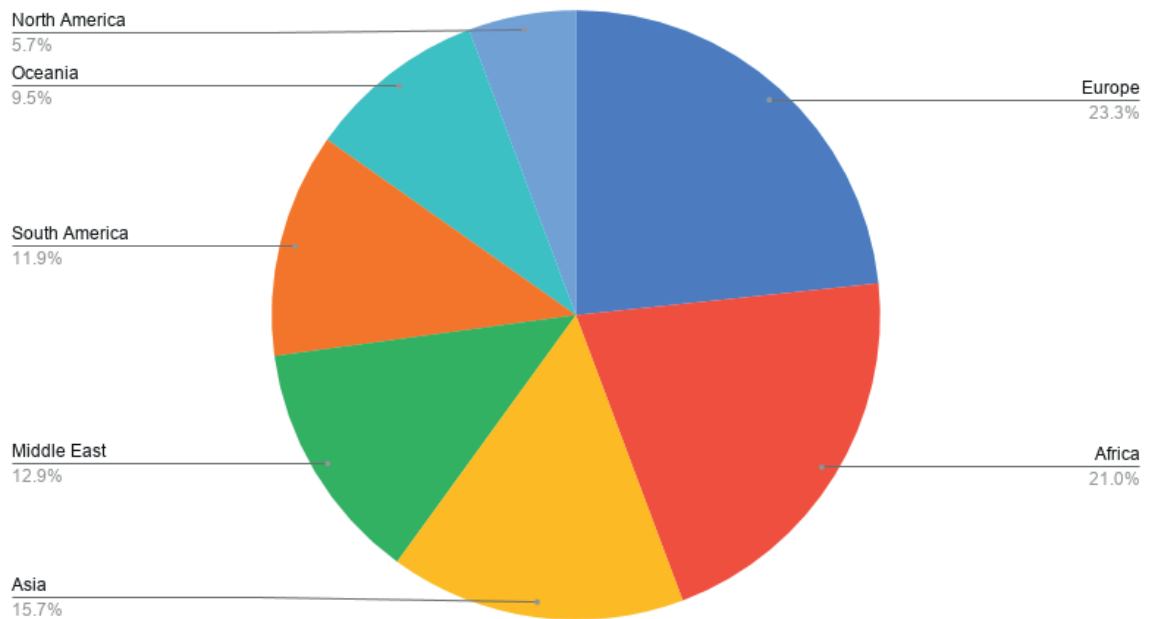


Figure 6. A view of Facebook third-party fact checking network by continent and operations

The Third Party Fact-Checking programme was initially launched in the US in mid-December 2016 with The Associated Press, Politifact, FactCheck.org, Snopes and ABC News. It then expanded rapidly internationally. In June 2018, three months after the Cambridge Analytica scandal broke, the Program linked 25 organisations in 14 countries. In early September 2020, when this research was completed, Facebook partnered with 74 third party fact checking organisations around the world, in over 50 languages (this data analysis is based on Facebook’s former list of partners and their newest partners’ map⁴⁵).

⁴⁵ https://web.archive.org/web/20200728165712if_/https://www.facebook.com/business/help/182222309230722 (deprecated by Facebook in August 2020 and replaced by a map: <https://www.facebook.com/journalismproject/programs/third-party-fact-checking/partner-map>)

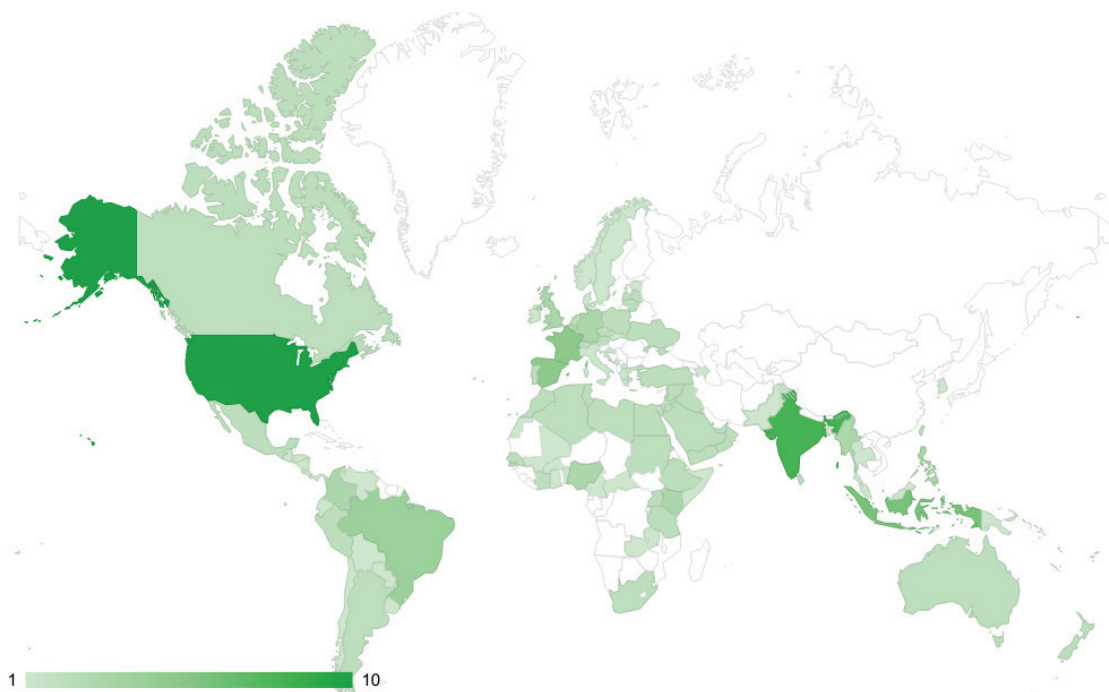


Figure 7. Map view of Facebook third-party fact checking network worldwide distribution

The above map (Figure. 7) shows the state of Facebook Third-Party Fact Checking programme in September 2020. Table. 2 below outlines the distribution by number of fact checking operations being contracted by Facebook as of 10 September 2020).

Coverage	Number of operations
United States of America	10
India	8
Indonesia	6
France, Spain	5
Brazil	4
Belgium; Columbia; Democratic Republic of Congo; Kenya; Myanmar; Nigeria; Philippines; Germany; United Kingdom;	3
Algeria; Argentina; Australia; Austria; Bahrain; Canada; Czech Republic; Egypt; Ethiopia; Iraq; Ivory Cost; Jordan; Kuwait; Latvia; Lebanon; Libya; Lithuania; Mexico; Morocco; Netherlands; New Zealand; Oman; Palestine; Peru; Poland; Portugal; Qatar; Saudi Arabia; Senegal; Singapore; South Africa; Sri Lanka; Sudan; Switzerland; Syria; United Republic of Tanzania; Tunisia; Turkey; Uganda; Ukraine; United Arab Emirates; Yemen	2
Azerbaijan; Bangladesh; Benin; Bolivia; Burkina Faso; Burundi; Cameroon; Central African Republic; Chile; Cook Islands; Costa Rica; Croatia; Denmark; Ecuador; El Salvador; Estonia; Fiji; French Polynesia; Ghana; Greece; Guatemala; Guinea-Conakry; Honduras; Ireland; Israel; Italy; Kiribati; Luxembourg; North Macedonia; Malaysia; Mali; Marshall Islands; Micronesia; Montenegro; Nauru; New Caledonia; Nicaragua; Niue; Norway; Pakistan; Palau; Panama; Papua New Guinea; Paraguay; Samoa; Slovakia; Solomon Islands; Somalia; Republic of Korea; Sweden; Thailand; Tonga; Tuvalu; Uruguay; Vanuatu; Venezuela; Zambia	1

Table 2. Distribution of Facebook's third-party fact checking network by number of operations

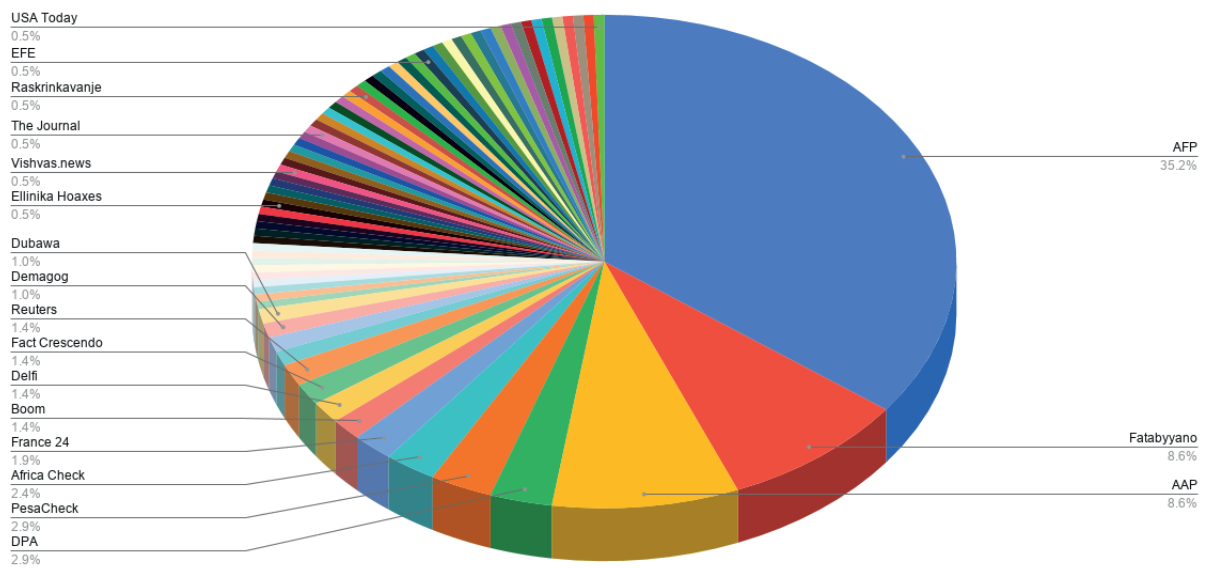


Figure 8. *Distribution of Facebook third-party fact checking programme by organisations involved*

Fact-checkers are selected and remunerated⁴⁶ by Facebook. As a prerequisite, all fact-checkers must be certified by the IFCN and adhere to their Code of Principles.

This programme represents the biggest organised international network dealing with disinformation, and it covers both Facebook and Instagram (since May 2019)⁴⁷, but it is relevant more widely, as false or misleading content on Facebook is often cross-posted on other social networks like Twitter, YouTube or (Facebook-owned) WhatsApp. So, curtailing the spread of disinformation on Facebook and Instagram can theoretically limit it ‘jumping’ to WhatsApp and beyond. Although WhatsApp does not directly send contested content to fact-checkers, it has a (little-publicised) chatbot which enables users to get tips on fact-checking and links to local checkers via the IFCN database who can be approached to investigate⁴⁸.

In response to COVID-19, IFCN also led the creation of a WhatsApp chatbot that lets users search IFCN’s dedicated database of COVID-19 fact-checks (Grau, 2020). In addition, as discussed further in chapter 7.3, in August 2020 WhatsApp started testing (in six countries) a new feature which allows users to carry out simple fact-checking of viral messages themselves, by searching the content on Google (Sweney, 2020). Regarding Instagram, posts rated false by third-party-fact checkers are removed from Instagram’s Explore and hashtag pages. In addition, content in Instagram feed and Stories that has been rated false by third-party-fact checkers is down-ranked.⁴⁹

The US in the run-up of the 2020 presidential election has become the top country with ten fact checking organisations, followed by India (with eight), where the spread of

⁴⁶ See discussion below about transparency issues regarding the fees involved.
⁴⁷ <https://about.fb.com/news/2019/12/combating-misinformation-on-instagram/>
⁴⁸ <https://faq.whatsapp.com/general/ifc-n-fact-checking-organizations-on-whatsapp>
⁴⁹ <https://about.instagram.com/blog/announcements/coronavirus-keeping-people-safe-informed-and-supported-on-instagram/>

disinformation through word-of-mouth or WhatsApp has fuelled mob violence reportedly resulting in deaths (McLaughlin, 2018), and Indonesia (with six).

The pie chart distribution by organisations in Fig. 8 clearly shows that Agence France-Presse (AFP) news agency has taken a leading share in the programme by launching fact checking operations in more than 70 countries with 90 journalists (including Argentina, Australia, Belgium, Bolivia, Brazil, Burkina Faso, Cameroon, Canada, Chile, Colombia, Costa Rica, Czech Republic, Ivory Coast, Democratic Republic of Congo, Ecuador, El Salvador, France, Ethiopia, Germany, Guatemala, Honduras, India, Indonesia, Kenya, Malaysia, Mexico, Netherlands, Nicaragua, several countries of Middle East and North Africa (mainly from Lebanon), Nigeria, Pakistan, Panama, Peru, Philippines, Poland, Senegal, Singapore, Slovakia, Somalia, South Africa, Spain, Sri Lanka, Switzerland, United Republic of Tanzania, Thailand, Tunisia, Uganda, United States of America, and Uruguay).

AFP made clear in December 2018 that AFP has made the fight against disinformation a core component of its mission, urging that other news agencies have an obligation to debunk false and manipulated stories (Fries, 2018).

Other well known mainstream media embracing fact checking and debunking within Facebook's Third Party Fact-checking programme include: The Associated Press (AP in the United States), The Australian Associated Press (in Oceania), Reuters, German Deutsche Press Agentur (DPA; in Germany, Belgium and Luxembourg), French international broadcaster France 24 Observers team (in 4 countries: Ivory Coast, Democratic Republic of Congo, Guinea-Conakry and France), Rappler (the Philippines), The Quint (India), and French daily *Libération* (France).

Despite the value placed on transparency in fact-checking processes outlined above, there is very limited transparency about how much Facebook pays its third-party fact-checking partners. In a report published in July 2019, British fact-checking operation Full Fact acknowledged that they received £171,800 (for 96 fact-checks) during the first six months of their involvement in the partnership (Hazard Owen, 2019). The money earned depends in part on the amount of fact checking done under the programme. French daily *Libération* and its fact checking service checknews.fr explained that they earned \$240,000 in 2018 for 249 articles uploaded to Facebook (Checknews, 2019). [Factcheck.org](http://factcheck.org) (U.S.) earned an amount of \$242,400 during fiscal year 2019 (12 months period ending on June 30, 2019) and \$59,500 in the next quarter (1st quarter of fiscal year 2020, ending on September 30, 2019) (Factcheck, 2019). The amount of debunked articles during those periods was not disclosed. Snopes (U.S.), while pulling out from the Facebook partnership in December 2018, disclosed having earned \$406,000 in 2018 and \$100,000 in 2017 (without reference to the number of debunked suspicious claims). According to fact-checkers' contracts described to the BBC, for each explanatory article, Facebook pays a fixed fee, which, in the U.S., is understood to be around \$800 (£600) (Lee, 2019b).

Evaluation of the Facebook initiative

[Poynter.org](http://poynter.org) conducted a survey of 19 organisations partnering with Facebook (Funke & Mantzarlis, 2018a) which identified a range of reasons underpinning their decision to join the Network. One of those partners, the British Full Fact referenced above, joined the Facebook Third Party Fact-checking programme late in 2018 and published an evaluative report (Full Fact, 2019) six months into the contract. It remains the most detailed evaluation of the functioning of the programme. While considering Facebook's Third Party Fact-Checking programme as "worthwhile" and likely "needed" for other internet communications companies too, Full Fact also raised some important issues and recommendations like the need for Facebook to:

1. "Fully include Instagram content" into the web interface providing a "queue" of suspicious content provided to fact checkers.⁵⁰
2. "Develop more tools to enable fact checkers to search for and surface similar content" to address "a repeated pattern with online misinformation" and avoid addressing only the tip of the iceberg.
3. Provide more data (to fact checkers) on "shares over time for flagged content"
4. "Share more data with fact checkers about the reach of our fact checks" in order to assess the value of the work undertaken within the program.

Full Fact also regards Facebook's internal rating system - false, mixture, false headline, true, not eligible, satire, opinion, prank generator, and not rated - as 'ill-suited' to the purpose of fact-checking. The first three labels are used by Facebook to reduce the distribution of content and to notify users that this content has been fact checked. Full Fact complained that the 'mixture' label was insufficient as well as over-punitive - it is applied when content is considered a mix of accurate and inaccurate information used for unproven claims and thus the content distribution is downplayed accordingly.

Reacting to Mark Zuckerberg's statement before the U.S. Congress, foreseeing an increasing shift towards a method where more of this content is flagged up front by Facebook A.I. (Artificial Intelligence) tools (Zuckerberg, 2018), Full Fact said it would welcome a clearer statement from the company about "the potential avenues they see for developing machine learning tools" based on the Third Party Fact Checking Partnership data.

Overall, according to the above-mentioned Poynter survey, judged by their own objectives, fact-checkers appear moderately satisfied with the Facebook partnership and the payment they receive for their work. The most critical question for these Facebook partners, according to the Poynter survey, remains concern that the company is not telling the public enough about how the partnership works. But the survey also demonstrated that there is also a lack of certainty about the efficacy of the initiative in terms of actually reducing disinformation on Facebook.

Investigations carried out by BuzzFeed concluded that "Facebook is still the home of viral fake news" (Silverman et al., 2017; Silverman & Pham, 2018). But there was some evidence of a reduction in engagement with disinformation on Facebook after the 2016 U.S. presidential election, with three studies concluding that this could be partially attributed to fact-checking interventions. A study from researchers at NYU and Stanford universities concluded that engagement (shares, likes, comments) with 'fake news' on Facebook fell from a peak of roughly 200 million per month at the end of 2016 to approximately 70 million per month in July 2018 (Allcott et al., 2018). The researchers noted that "...efforts by Facebook following the 2016 election to limit the diffusion of misinformation *may* have had a meaningful impact." The ratio of misinformation and disinformation detected on both Facebook and Twitter also "declined sharply" according to the study, "...from around 45:1 during the U.S. 2016 election to around 15:1 two years later." Nevertheless, according to this research, Facebook remains a much bigger disinformation vector than Twitter.

Another academic study from the University of Michigan introduced an "Iffy Quotient" to describe websites that frequently publish misinformation (Resnick et al., 2019). The study concluded that Facebook and Twitter did a poor job during the 2016 election season,

⁵⁰ Since Full Fact's report was published, Instagram content is now subject to fact-checking too, as noted above.

with the distribution of information from questionable sites doubling compared to the rate earlier that year. “However, there has been a long-term decline in Facebook’s ‘Iffy Quotient’ since March 2017”, the authors noted.

In further research, Décodeurs, the fact-checking operation of French daily *Le Monde*, analysed 630 French websites in 2018 with the help of their Decodex browser extension which warns web surfers if they reach a dubious news website or a known disinformation source of another kind. They concluded that engagement with low accuracy and dubious websites as well as virality of false news decreased significantly on Facebook (Sénécat, 2018).

Facebook highlighted these studies in a 2018 blog post stating that they represented evidence that the “...overall volume of false news on Facebook is trending downward” (Lyons, 2018b).

More recently, an announcement from Facebook sparked a controversy about the company’s policy regarding the fact-checking of political advertising. The company had decided that it would not send organic content or adverts from politicians or their affiliates to its third-party fact-checking partners for review (Clegg, 2019).

Early 2019, a few months after ABC News (U.S.) dropped out of the Facebook fact-checking programme, the anti-hoax U.S. website Snopes decided to quit the Facebook Third Party Fact-Checking programme despite earning 33% of its income in 2018 from the partnership (Green & Mikkelson, 2019). At the end of November 2019, Dutch fact-checker [Nu.nl](https://www.nu.nl/) announced their withdrawal from the programme amid controversy around the exemption of certain categories of political advertising (see below) from fact-checking by partners (Hern, 2019a).⁵¹

Facebook’s policy generally exempts political speech from fact-checking, in the form of posts and adverts made by politicians, political parties and affiliates. However, the policy provides that fact-checking can cover “organisations such as Super PACs or advocacy organisations that are unaffiliated with candidates”. It also states that:

“*When a politician shares a specific piece of content - i.e., a link to an article, video or photo created by someone else that has been previously debunked on Facebook - we will demote that content, display a warning and reject its inclusion in ads. This is different from a politician’s own claim or statement. If a claim is made directly by a politician on their Page, in an ad or on their website, it is considered direct speech and ineligible for our third party fact checking program – even if the substance of that claim has been debunked elsewhere.*”⁵²

However, as this study was being finalised in July 2020, Facebook removed a piece of content posted by President Trump for the first time, for violating its COVID-19 related policies. The post included a clip of him claiming that children were “almost immune” to coronavirus. According to a company spokesperson: “This video includes false claims that a group of people is immune from COVID-19 which is a violation of our policies around harmful COVID misinformation.” (BBC, 2020d; Carrie Wong 2020)

⁵¹ See also chapter 7.1

⁵² <https://www.facebook.com/business/help/182222309230722>

Regarding editorial independence, in late 2018 Politifact issued a statement on Twitter endorsed by [Factcheck.org](https://factcheck.org) (U.S.), Agência Lupa (Brazil) and [Teyit.org](https://teyit.org) (Turkey) to deny a report from *The Guardian* claiming that “Facebook pushed reporters to prioritise the debunking of misinformation that affected Facebook advertisers” (PolitiFact, 2018).

According to some news organisations undertaking debunking as members of the programme, Facebook does not prevent them from fact-checking content from politicians and political parties (including political advertising) but they do not pay them to undertake this work and this content is not labelled on the platform when found to be false or misleading by the fact-checkers. For instance, in 2019, AFP fact-checked statements from far-right French leader Marine Le Pen five times on its Factual blog (Daudin, 2019) and on its Facebook account, but this was not reflected within the Facebook ecosystem due to its policy limiting the fact-checking of much political content.

Opinion content is another contested area. Facebook policy states that opinion is “generally not eligible to be rated by fact-checkers. This includes content that advocates for ideas and draws conclusions based on the interpretation of facts and data, and tells the public what the author or contributor thinks about an event or issue.”⁵³ The policy includes a caveat that “...content presented as opinion but based on underlying false information may still be eligible for a rating.” However, this policy has loopholes that have resulted in criticism and controversy. One example is a case pertaining to an op-ed from a climate change denialist group which was based on false and misleading assertions about climate science. In this case, Facebook’s climate science fact-checking partner Climate Feedback rated the article as “false”⁵⁴, however following an appeal from the lobby group, Facebook removed the label on the basis that the article was an “opinion” and ineligible for fact-checking (Penney, 2020; Pasternak 2020). In another example, a “false” label applied by medical fact-checkers to a video published on an anti-abortion activist’s Facebook page claiming that abortion was never medically necessary was removed by Facebook following multiple complaints from conservative lawmakers (Grossman & Schickler, 2019). Although the International Fact Checking Network investigated the fact-checkers’ determination and found in September 2019 that the video claim was indeed false⁵⁵, the video was still proliferating on Facebook a year later with no fact-checking label⁵⁶.

This last example, in particular, prompted lawmakers in the UK House of Lords to note in their report from the inquiry into Digital Technology and the Resurrection of Trust that: “There were no material concerns with the accuracy of the fact check and it was supported by an independent secondary review... This suggests that Facebook’s position is more about avoiding political pressure than any particular concern about preserving democratic debate.” (House of Lords, 2020).

The Facebook political advertising controversy (concerning its policy on fact-checking noted above) will be covered further in chapter 5.3 on electoral-specific responses, and chapter 7.1, which focuses on ethical and normative responses.

⁵³ https://www.facebook.com/business/help/315131736305613?recommended_by=297022994952764

⁵⁴ <https://climatefeedback.org/evaluation/article-by-michael-shellenberger-mixes-accurate-and-inaccurate-claims-in-support-of-a-misleading-and-overly-simplistic-argumentation-about-climate-change/>

⁵⁵ <https://www.poynter.org/fact-checking/2019/the-ifcn-concludes-investigation-about-science-feedback/>

⁵⁶ <https://www.facebook.com/youngamericasfoundation/videos/2113086642330235>; <https://www.facebook.com/youngamericasfoundation/videos/2113086642330235>

b. Regional responses

AfricaCheck

AfricaCheck⁵⁷ has been the main driver of fact-checking in Africa. It is a non-profit organisation set up in 2012 to promote accuracy in public sphere debate and within the news media in Africa. The goal was to raise the quality of information available to society across the continent. Devised initially by the AFP Foundation, a non-profit media development arm of the international news agency AFP, Africa Check is an independent organisation with offices in Johannesburg (South Africa), Nairobi (Kenya), Lagos (Nigeria) and Dakar (Senegal).

It produces reports in English and French, testing claims made by public figures, institutions and the media against the best available evidence. Since 2012, it has fact-checked more than 1,500 claims on topics from crime and race in South Africa, to population numbers in Nigeria, and fake health 'cures' in various African countries.

Africa Check's work is published and discussed in the news media across the continent. Its head office is based at the Journalism Department of the University of the Witwatersrand in Johannesburg, South Africa, while the French language site has been run by a team based at the EJICOM journalism school in Dakar, Senegal, since 2015. Africa Check relies on its readers to identify the claims they want checked, and it also enables and encourages other journalists to check claims themselves with the assistance of a fact-check section, research reports and teaching services.

Since its creation, Africa Check has received funding support from the AFP Foundation, the Aga Khan University, the Journalism Department of the University of the Witwatersrand, and the EJICOM journalism school, as well as grants from a long list of philanthropic institutions including The African Media Initiative and African News Innovation Challenge, The International Press Institute (IPI), Google, the Konrad Adenauer Stiftung, the Millennium Trust, Luminare, the Open Society Foundations (OSF), the Shuttleworth Foundation, the Bill and Melinda Gates Foundation, the Raith Foundation, Standard Bank, Absa and Code for Africa.

The Duke University fact-checking database registered 17 active fact-checking organisations in Africa in early 2020.

Latin America: the influence of Chequeado

In Latin America, Argentina's Chequeado⁵⁸ has been prominent in the fact-checking community since its creation in 2010. Many new initiatives have emerged in the region since 2014, mostly in the journalism field, thanks to the help and influence of Chequeado, especially in the area of fact-checking methodologies. In 2019, Chequeado coordinated with AFP on the Reverso project to fact-check the Argentinian presidential election campaign.

Duke University's fact-checking database registers 16 organisations in South America, eight in Central America in Spanish, and 10 in Portuguese (in Europe, there are 6 organisations in Spain and 2 in Portugal). In 2014, Chequeado invited the other regional fact-checking organisations to a meeting in Buenos Aires to launch a new network "LatamChequea" designed to exchange best practices. Since then, the regional network

⁵⁷ <https://africacheck.org/>

⁵⁸ <https://chequeado.com/>

has been holding a biannual conference in Buenos Aires. There are also monthly virtual meetings between fact-checkers which also involve a number of social scientists.

Chequeado is supported financially by a foundation, La Voz Pública, and it is active in research collaborations with academics. Fact-checkers are also embedded for a week or two in Chequado's newsroom with the support of IFCN scholarships.

Europe: SOMA

The Social Observatory for Disinformation and Social Media Analysis (SOMA) is funded by the European Commission with the objective of organising European fact-checkers as part of a pan-European effort to rebuild trust in journalism, and to provide support to the growing community of media organisations, fact-checkers, academics, and NGOs and policy makers fighting disinformation on the continent. In the first year of operation, some 40 European organisations have formally join this Observatory, based on the platform [Truly Media](#)⁵⁹. This European Observatory has published several investigations and recommendations regarding disinformation around the COVID-19 pandemic. This observatory is due to be continued in the forthcoming years by a new one called EDMO (European Digital Media Observatory).⁶⁰

Arab States

In the Arab countries, collaboration between fact-checking initiatives is not institutionalised but fact-checkers in the region are connected, collaborate on training, and gather at conferences such as Alexandria Media Forum in Egypt which has focused on fact-checking, disinformation, and media literacy and training in its three last editions in Alexandria (2017-2018) and Cairo (2019).

Regionally, one prominent initiative, launched in 2014, is Jordan-based [Fatabyyano](#)⁶¹. Launched in 2014, it monitors and debunks disinformation in eighteen countries in the Middle East and North Africa. Others include [Da Begad](#)⁶² ('Is it real?') launched in Egypt in 2013, as well as [Matsad2sh](#)⁶³ ('Don't believe') and [Falsoo](#)⁶⁴. Homonyms Falso work on fact-checking in [Libya](#)⁶⁵ and in [Tunisia](#)⁶⁶.

In the Syrian Arab Republic, [Verify Syria](#)⁶⁷ is publishing a monitoring and debunking website in three languages, Arabic, English and Turkish. The [AFP fact-checking operation](#)⁶⁸ covering various countries of Middle East and North Africa is based in Lebanon as a collaboration with the local fact-checker [El3asas](#)⁶⁹.

c. Some other national responses

This subsection details specific and noteworthy national initiatives in the area of monitoring and fact-checking. In the U.S. and in Europe, the history and evolution of fact-

59 <https://www.truly.media/>

60 <https://edmo.eu/>

61 <https://fatabyyano.net/> and <https://www.facebook.com/Fatabyyano/>

62 <https://dabegad.com/>

63 <https://www.facebook.com/matsda2sh/>

64 <https://www.falsoo.com/>

65 <https://falso.ly/>

66 <https://www.facebook.com/falso.tn/>

67 <https://www.verify-sy.com/>

68 <https://factual.afp.com/ar>

69 <https://twitter.com/el3asas>

checking is tied to election campaigns and verifying political claims. Therefore most of these responses are analysed in chapter 5.3.

India

In India, Facebook-owned WhatsApp has developed into one of the main channels of disinformation (Kajimoto & Stanley, 2019). It has been noted that the phenomenon of politicians using social media to directly access audiences, bypassing traditional media gatekeepers, has been identifiable since 2014 and it has aided the spread of disinformation within online social networks (Kaur & Nair, 2018).

Fifteen active fact-checking organisations operate in India according to Duke University's database and eight are members of Facebook's third-party fact-checking network. Nevertheless, those outlets are mostly individuals, and small organisations or teams (like the Times of India fact-checkers). All of these have been created since 2013. They include [Factcheck.in](https://www.factcheck.in/)⁷⁰, [SM Hoax Slayer](https://smhoaxslayer.com/)⁷¹, and investigative journalism outlet Boomlive, which pivoted to fact-checking in 2016. A large part of the disinformation they debunk is political, either local or about geopolitical tensions.

Indonesia

In Indonesia, the NGO Mafindo has been fighting disinformation since 2015 through an [anti-defamation and hoax group](https://www.facebook.com/groups/fafhh)⁷² on Facebook, a [WhatsApp-based hoax buster](https://mafindo.gitbook.io/whatsapp-hoax-buster/)⁷³, a Google Chrome extension, and a website⁷⁴ using their motto "Turn Back Hoax". Following Mafindo, five other debunking initiatives have been launched in Indonesia, mostly by news organisations. Six of them are part of the Facebook third-party fact-checking network. Another initiative mentioned by researchers (Kajimoto & Stanley, 2019), Klarifikasihoax has been inactive since 2017.

Philippines

In the Philippines, disinformation campaigns are collectively creating a public sphere filled with information pollution and, consequently, with toxic incivility and polarisation since the 2016 presidential election. As reported in a UNESCO publication (Posetti 2017), troll armies using 'sock puppet' networks have gained traction with potentially long-term consequences for democracy and elections (see also Ong & Cabañes, 2018; Ressa, 2016). However, four fact-checking organisations are monitoring disinformation and political claims, including [Vera files](https://verafiles.org/)⁷⁵ and [Rappler](https://www.rappler.com/newsbreak/fact-check)⁷⁶, and three of them are members of the Facebook third-party fact-checking network.

Republic of Korea

In the Republic of Korea, there has been a proliferation of rumours, partisan propaganda and disinformation on mobile messaging apps like KakaoTalk or Naver Band, as well as social media sites, especially during elections. One of the main initiatives set up for the 2017 presidential election was [SNU Factcheck](http://factcheck.snu.ac.kr/)⁷⁷, launched by Seoul National University to gather 26 news outlets to cross-check disputed information. It continues as one of the

⁷⁰ <https://www.factchecker.in/about-us/>

⁷¹ <https://smhoaxslayer.com/about/>

⁷² <https://www.facebook.com/groups/fafhh>

⁷³ <https://mafindo.gitbook.io/whatsapp-hoax-buster/>

⁷⁴ <https://turnbackhoax.id/>

⁷⁵ <https://verafiles.org/>

⁷⁶ <https://www.rappler.com/newsbreak/fact-check>

⁷⁷ <http://factcheck.snu.ac.kr/>

five fact-checking organisations in the country. The Facebook third-party fact-checking network does not have a Korean member, mainly because the local Naver and Daum are the most popular online portals with a policy of asking news content providers to go through an evaluation process and thereby making it harder for disinformation purveyors to syndicate content through those portals.

U.S. - Snopes

Snopes is one of the debunking and fact-checking pioneers in the U.S.. Back in 1994, founder David Mikkelson created [snopes.com](https://www.snopes.com/)⁷⁸ as a hobby to investigate urban legends and hoaxes on the Usenet (a worldwide discussion channel of the early internet) newsgroup alt.folklore.urban.

Immediately after the 9/11 terrorist attacks in 2001, the founders of [snopes.com](https://www.snopes.com/) started to debunk rumours and lies about the attacks - a total of 176 legends and rumours⁷⁹ were evaluated by Snopes between 2001 and 2011 (Aspray & Cortada, 2019). This was the inflection point for [snopes.com](https://www.snopes.com/) to shift from demystifying urban legends as a hobbyist, to progressively becoming a major fact-checking organisation (Dean, 2017). Between breaking news (like Hurricane Katrina in 2005) and presidential elections (2008 with the rumours circulating about Barack Obama's place of birth; up to the 2016 poll), [snopes.com](https://www.snopes.com/) grew its audience, allowing it to build a sustainable business through advertising revenue.⁸⁰

4.1.5 Response Case Study: COVID-19 Disinformation

WHO Director-General Tedros Adhanom Ghebreyesus has stated that the Coronavirus outbreak has left humanity not just fighting an epidemic but also "an infodemic" (Zarocostas, 2020). All organisations reviewed in this chapter have taken measures to respond to the COVID-19 crisis with special hubs or pages about COVID-19 disinformation.

For example, First Draft has published a whole hub⁸¹ of resources for reporters such as tools, guides, ethics guidelines, an online course, and a searchable database of coronavirus debunks based on two monitoring tools: [Google Fact Check Explorer](https://toolbox.google.com/factcheck/explorer)⁸², and the [IFCN CoronaVirusFacts Alliance database](https://www.poynter.org/ifcn-covid-19-misinformation/)⁸³. The latter was launched in January 2020 as a double hashtag campaign on Twitter #CoronaVirusFacts (in English) and #DatosCoronaVirus (in Spanish) for participating IFCN members, when the epidemic was still limited to China but was already being exploited for disinformation purposes.

The hashtags campaign led to a database of more than 3000 fact-checks from 70 countries and 40 languages (in April 2020). Then, another project led by [Science Feedback](https://sciencefeedback.co/)⁸⁴, and sponsored by the Google News Initiative, sought to expand this database with all the urls sharing COVID-19 disinformation.

⁷⁸ <https://www.snopes.com/>

⁷⁹ https://scholar.colorado.edu/concern/book_chapters/8049g572m

⁸⁰ See the earlier discussion in this chapter, and in chapter 7.1, Snopes' role as a member of the Facebook Third Party Fact-checking Network

⁸¹ <https://firstdraftnews.org/long-form-article/coronavirus-resources-for-reporters/>

⁸² <https://toolbox.google.com/factcheck/explorer>

⁸³ <https://www.poynter.org/ifcn-covid-19-misinformation/>

⁸⁴ <https://sciencefeedback.co/building-an-open-source-database-of-misinformation-sources-on-covid-19/>

The fact-checking community, from the IFCN, Facebook's Third Party Fact-Checking programme and beyond, have published countless debunking reports about the Coronavirus outbreak, registering disinformation cases from all continents. According to a study from the Reuters Institute for the Study of Journalism (RISJ) based upon an analysis of English-language fact-checks curated by First Draft, the number of fact-checks increased more than 900% from January to March 2020. On the 225 debunks analysed, RISJ found that 59% of the misinformation content was reconfigured while 38% was fabricated. (Brennan et al 2020)

Some internet communications companies (e.g. Facebook⁸⁵, YouTube⁸⁶, Instagram⁸⁷, WhatsApp⁸⁸, Twitter⁸⁹, LinkedIn⁹⁰) themselves have taken action to connect their users to reliable information about the pandemic by linking any query on the coronavirus to the World Health Organisation (WHO) main hub⁹¹ and their WHO mythbusters page⁹², or to the local government's ministry of health. They are also relaying alerts from WHO through chatbots, and from local authorities on message applications⁹³ too, or publishing curated official pages of factual information. Some are also promoting the IFCN affiliated fact-checking organisations⁹⁴ and asking their users to verify the facts and to refrain from sharing information if they are not sure it is true.

Google (Mantzaris, 2020), Facebook (Goldshlager & Watson, 2020) and WhatsApp (IFCN, 2020a) announced small programmes to fund fact-checkers and nonprofits fighting disinformation about the pandemic in several countries (IFCN, 2020b). Thirteen projects in the same number of countries were announced at the beginning of April through the "Coronavirus Fact-Checking Grants" (IFCN, 2020c) program.

In addition, internet communications companies have decided to "work closely together"⁹⁵ to combat fraud and misinformation connected to the pandemic. Many companies have started blocking adverts that try to capitalise on coronavirus-related disinformation and removing disinformation that could lead to physical harm. For example, in April Facebook said that it put 50 million warning labels on pieces of content on the platform, based on over 7,500 articles from their fact-checking partners.⁹⁶ Some are also removing conspiracy-type content, using policy provisions about content consideration in terms of its likely potential to cause harm.⁹⁷ In Facebook's case, this provision is "Misinformation and unverifiable rumors that contribute to the risk of imminent violence or physical harm."⁹⁸ Facebook CEO Zuckerberg stated that it was "easier" to make the difference between good and wrong information in a pandemic than in a political campaign (Smith, 2020a).

It is not possible to accurately gauge the extent of fact-checked COVID-19 disinformation within the companies, because they typically do not provide access to granular statistics

⁸⁵ https://www.facebook.com/coronavirus_info/?page_source=coronavirus_hub_attachment&fref=mentions

⁸⁶ https://www.youtube.com/watch?v=i352PxWf_3M

⁸⁷ <https://about.instagram.com/blog/announcements/coronavirus-keeping-people-safe-informed-and-supported-on-instagram/>

⁸⁸ <https://www.whatsapp.com/coronavirus>

⁸⁹ https://blog.twitter.com/en_us/topics/company/2020/covid-19.html

⁹⁰ <https://www.linkedin.com/feed/news/coronavirus-official-updates-4513283/>

⁹¹ <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>

⁹² <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public/myth-busters>

⁹³ <https://www.messenger.com/coronavirus>

⁹⁴ <https://faq.whatsapp.com/126787958113983>

⁹⁵ <https://twitter.com/googlepubpolicy/status/1239706347769389056>

⁹⁶ <https://about.fb.com/news/2020/04/covid-19-misinfo-update/>

⁹⁷ https://blog.twitter.com/en_us/topics/company/2020/covid-19.html

⁹⁸ https://www.facebook.com/communitystandards/credible_violence/

about the origins, numbers and types of items checked, nor data on the circulation of such content prior to and post any actions being taken. (UNESCO, 2020)

Notwithstanding fact-checking efforts, briefing papers issued by the Institute for Strategic Dialogue (ISD, 2020a) in March and April 2020 have warned about the exploitation of COVID-19 pandemic by anti-migrants and xenophobic or far-right networks (ISD, 2020b), especially in closed groups on Facebook, chat channels on WhatsApp, and fringe networks, like 4chan (Arthur, 2019), as well as in languages other than English.

4.1.6 How are monitoring and fact-checking responses evaluated?

Fact-checking can be evaluated in terms of whether it is achieving its immediate and longer term objectives. This depends on assessing its volume, reach and timeliness, among other factors. However, this is not straightforward as there is still limited published research on the reach and impact of fact-checking. Much of the relevant data is held in private by the internet companies. This makes evaluation difficult, and leaves researchers to make extrapolations from limited data (such as reach and engagement metrics attached to debunks and fact-checks published by news organisations), and audience research (e.g. ethnographic, psychological studies) into diverse citizens' responses to both disinformation and corrective measures. Further, the underlying assumption that verified evidence and rational thought have a role to play in countering disinformation is hard to test empirically because of the complex interlinkages of disinformation with emotion and identity.

During the run-up to the 2012 U.S. presidential election, concerns arose about the efficiency of fact-checking and its ability to reduce disinformation, particularly that connected to political rhetoric. But there was still sufficient momentum for the continuation of fact-checking efforts, as evident in the words of one researcher: "Naming and shaming politicians who repeatedly mislead the public can still inflict significant reputational damage over time" (Nyhan, 2012).

Promoting more involvement of citizens in public affairs, increasing politicians' reputational cost, and increasing the public's trust in the news media have been identified in several studies as having positive effects for fact-checking. However, the predisposition of citizens to accept corrections that reinforce their own views is relevant. The backfire or 'boomerang effect' helps to spread disinformation (i.e. if fact-checks contradict citizens' pre-existing views about a political actor or issue, they are more likely to be rejected despite their accuracy), especially when disinformation (and fact-checks) are weaponised by the politicians themselves to increase polarisation.

One fact-checking organisation that has tried to assess the impact of its work is Argentina's Chequeado. In a review of six academic studies assessing the impact of fact-checking in the United States, researchers commissioned by Chequeado to study the efficacy of their efforts considered the impact of fact-checking on citizenry, political elites, and media organisations. (Pomares & Guzman, 2015) They found that promoting the involvement of citizens in public affairs, increasing the reputational cost of spreading falsehoods for politicians, and aiding public trust in the news media are positive effects of fact-checking. One evaluative response proposed by the Chequeado-commissioned researchers is to measure the strength of the fact-check (e.g. how well did it stand up to scrutiny?) and its reach in tandem. (Pomares & Guzman, 2015)

A recent joint report from Africa Check, Chequeado and Full Fact listed public corrections of misleading statements or statistics, stopping false claims from politicians, releasing new meaningful data, getting journalists in legacy newsrooms trained to reduce the spread of disinformation, and engaging with officials, and efforts to raise accountability as potential benefits of fact-checking (Africa Check, Chequeado & Full Fact, 2020). These could be treated as indicators for efficacy.

4.1.7 Challenges and opportunities

The volume and range of types of disinformation make it difficult to identify, monitor, report and draw public attention to all instances and all dimensions of the problem. There are also key nuances, such as it is one thing to demonstrate that a claim is false, another to show that it is currently without evidence (but potentially could be true), and a third to say that whether a particular proposition is factual when there is usually a wider narrative or perspective at play which mobilises and combines particular facts, as well as presents them along with opinion, attitude and identity.

This is further complicated by the task of assessment of the intended and unintended effects of identification of the content at hand, and of its providers. However, producing such analysis is vital in order to develop or modify fact-checking and other responses.

The challenge for fact-checkers is to aspire to objective standards and operate transparently in all countries and languages, at scale, and with impact. This is necessary to enable society to access the information required to ensure that the various responses are optimally effective. Achieving this in practice, however, is far from straightforward, especially in the case of non-global languages and smaller countries, which often do not have their own local independent fact-checking organisations. Instead, international fact-checking organizations aim to fill the gap, but inevitably need to rely on native speakers, limiting the possibility of scrutinising their work and biases. This highlights the need for a robust, independent approach to 'evaluating the evaluators' or 'fact-checking the fact-checkers.'

Fact-checking also needs to be consistent with international standards for freedom of expression and other human rights like privacy, and to recognise that certain content (e.g. unknowns, certain narratives, opinions, humour) does not lend itself to verification per se. Further, fact-checking has to live up to values of transparency and non-partisanship, and avoid selective instrumentalisation.

A challenge for fact-checking organisations is to fend off legal attacks on them. The Fact-checkers Legal Support Initiative has come into existence to assist with legal advice⁹⁹. It is a consortium of the Media Legal Defence Initiative, the International Fact-Checking Network and the Reporters Committee for Freedom of the Press.

Major events, such as elections and public health emergencies, provide an opportunity for independent monitoring and identification responses to reaffirm the value of facts, and to encourage public reflection of what content they treat as credible, and what people decide to share. For example, identifying COVID-19 disinformation and investigation of responses over time also enables continuous assessment of the internet communications companies' efficacy in "flattening the curve" of the 'disinfodemic' (Proctor, 2020; Posetti & Bontcheva 2020a; Posetti & Bontcheva 2020b). Identification responses are also key

⁹⁹ <https://factcheckerlegalsupport.org/>

for monitoring the intersection of disinformation with hate speech used against women, minorities, migrants and other vulnerable citizens and communities. However, it is acknowledged that some of these targeted groups may also resort to disinformation tactics and content produced by them should also be scrutinised. It is the case that sometimes groups that are victims of disinformation may themselves resort to the same tactics to further their various causes, and that their content should also be subject to fact-checking and scrutiny.

This is also an opportunity to strengthen identification responses. While WhatsApp (IFCN, 2020a), Facebook (Axelrod, 2020), [Google](#)¹⁰⁰, and Twitter (Gadde, 2020) have pledged some funding to fact-checking organisations, this also shows that more can be done. Ongoing support throughout and beyond critical periods of elections and pandemics is needed. Verifying claims about vaccinations and climate change is particularly significant going ahead.

4.1.8 Recommendations for monitoring and fact-checking responses

The challenges and opportunities identified above, and the current state of fact-checking and debunking, lead to the following policy recommendations for international and regional institutions, governments, internet communications companies, foundations and news organisations, which could:

- Make available resources for independent fact checking, including facilitating the fact-checking of political content and political advertising.
- Support the principle of access to information, especially in regard to both authorities and internet communications companies, as relevant to fact-checking, in order to increase transparency and enable fact-checking organisations themselves to work more accurately and transparently.
- Promote fact-checking results as trustworthy sources of information, useful for citizenship, for the news media, and for Media and Information Literacy interventions.
- Promote trans-disciplinary research into fact-checking responses to disinformation.
- Help to develop collaborative fact-checking operations worldwide to aid access to accurate and reliable information globally, especially in partnership with news organisations.
- Reinforce fact-checking capacity within news organisations through specialist training and editorial projects to support accountability reporting applied to corporate, government, and political actors and actions.
- ‘Verify the verifiers’ and develop international standards and an accountability approach to enable transparent, and objective appointment and assessment procedures for the people and organisations (including the internet communications companies that facilitate and fund fact-checking on their sites) involved in fact checking, and evaluate their performance over time.

¹⁰⁰ <https://www.france24.com/en/20200402-google-boosts-support-for-checking-coronavirus-facts>

4.2 Investigative responses

Authors: Sam Gregory, Julie Posetti and Denis Teyssou

This chapter addresses the range of entities producing investigations into disinformation (ranging from journalism to civil society investigations) and their outputs.

Investigative reports typically address particular campaigns and actors, and go beyond the questions of whether particular content contains falsehoods and the extent of the falsity (fact-checking). They may include, but also go beyond, the issue of whether a piece of content is manipulated or miscontextualised/misrepresented (such as in the case of provenance labelling), or whether a piece of content or outlet is reliable and fair (for example, credibility labelling). They are likely to monitor (as far as possible) the instances, themes and spread of the particular disinformation under focus. When fully deployed, they also provide insights into the dynamics of disinformation campaigns, including such elements as the networks conducting them, the targets, the mediums used, the methods used, budgets available, along with attribution and intent. For examples of categorisations and cataloguing of campaigns, see Bradshaw & Howard (2018) and Brooking et al. (2020).

Such investigations typically aim to help news organisations, governments, fact-checkers, Internet communications companies and others understand these dynamics, in order to deploy effective counter-measures to particular campaigns. They seek to alert actors to ongoing innovations in disinformation tactics and strategies. Increasingly, methodologies of categorisation are being developed to better catalogue across [related incidents](#).¹⁰¹

4.2.1 What and who do they target?

Investigative responses monitor a range of actors. A significant number focus on government-funded or -supported disinformation campaigns. Although many commercial and company responses initially began with a focus on one or two governments' roles in targeted disinformation campaigns, the range of state sponsors has expanded (Nimmo, 2019; Francois et al., 2019; Nimmo et al., 2019a; Gleicher, 2019a; Gleicher, 2019b). The '2019 Global Inventory of Organised Social Media Manipulation' report (Bradshaw & Howard, 2019) identifies government-implicated social media manipulation campaigns against foreign countries conducted by a number of States, while noting over 26 countries with internal disinformation activities. However, the entities above tend to have a blind-spot in regard to the covert or overt disinformational activities by certain governments in foreign countries or domestically. There is a focus on disinformation promoted by unofficial actors such as on white supremacist groups in the U.S. or far-right movements in India (the Southern Poverty Law Center, Equality Labs -see Soundararajan et al., 2019).

¹⁰¹ <https://www.bellingcat.com/>

Other investigators look at commercially-motivated or hybrid actors (albeit often working for political purposes). For example, foreign 'click-farms' engaged in significant disinformation interventions in U.S. politics for commercial reasons (Silverman & Alexander, 2016; Soares, 2017). Another example is investigation into the Epoch Media Group, a commercial entity with political motivations, which led a disinformation campaign including fake profiles and accounts (Nimmo, et al., 2019b). This was exposed via reporting from Snopes and the Operation #FFS investigation by [Graphika](https://graphika.com/)¹⁰²/ Facebook. An important trend in the past 2-3 years has been the growth of private sector disinformation-for-hire actors, providing services to politicians and political parties - as documented in a number of countries (Ong & Cabañes, 2019; Silverman et al., 2020). A 2019 survey by BuzzFeed News based on account takedowns by platforms, as well as publicised investigations by security and research firms "found that since 2011, at least 27 online information operations¹⁰³ have been partially or wholly attributed to PR or marketing firms. Of those, 19 occurred in 2019 alone." (Silverman, et al., 2020). Another important investigation was undertaken by South Africa's Daily Maverick into the now defunct UK-based PR firm Bell Pottinger which was exposed for artificially seeding racial tension in the country amid a state capture scandal linked to the presidency (Thamm 2019; Posetti et al 2019a).

4.2.2 Who do investigative responses try to help?

Investigative reporting serves a range of actors including companies engaged in detection of coordinated inauthentic behaviour on their platforms as well as official inquiries. One such inquiry was the U.S. Congressional investigation into foreign interference before and during the 2016 U.S. elections (U.S. Senate Select Committee on Intelligence, 2018), and another was the UK-initiated International Grand Committee on Disinformation and 'Fake News'¹⁰⁴. Governments also use these investigations, for example EU policy proposals based on commissioned reports (EU Disinfo Lab, 2019b). Coordination between internet communications companies and external actors investigating campaigns is often reflected in funding relationships between them (see below and chapters 4.1, 7.1).

Civil society organisations like Amnesty International also undertake forensic investigative work designed to detect, debunk and deter disinformation connected to human rights abuses. These include [Amnesty's Digital Verification Corps](https://www.theengineerroom.org/digital-verification-corps/)¹⁰⁵ - a partnership with six international universities that also collaborates on open source journalistic investigations (Fortune, 2018). Other stakeholders include individual citizens and the growing number of participants in the global anti-disinformation community. One such example is Amnesty International's 'Amnesty Decoders'¹⁰⁶ project, which crowdsources volunteer verification assistance to examine claims of human rights violations. Campaigning organisation Avaaz has also done investigations, including into the responses by internet communications companies, and advocated for changes accordingly.¹⁰⁷

¹⁰² <https://graphika.com/>

¹⁰³ https://www.militaryfactory.com/dictionary/military-terms-defined.asp?term_id=2637

¹⁰⁴ <https://www.oireachtas.ie/en/press-centre/press-releases/20191025-international-grand-committee-on-disinformation-and-fake-news-dublin-ireland-wednesday-6th-and-thursday-7th-november-2019/>

¹⁰⁵ <https://www.theengineerroom.org/digital-verification-corps/>

¹⁰⁶ <https://decoders.amnesty.org/>

¹⁰⁷ https://secure.avaaz.org/campaign/en/disinfo_hub/

The motivation behind investigative responses is to improve understanding of specific disinformation acts and campaigns so as to enable action to be taken against them. Such action could include content takedowns or demotion, legal processes, transparency and accountability measures, and regulatory or company policy reform. Investigative responses also aim to expose methods adopted in disinformation campaigns to impede further utilisation of these by malicious actors, and ensure knowledge on them is available to a burgeoning community of disinformation researchers. They identify structural challenges in disinformation as opposed to symptomatic examples or individual content items. As an example, the [EU Disinfo Lab](https://www.disinfo.eu/) initiative notes its commitment to “continuously monitor disinformation activities across the major platforms (digital and traditional), identify trends and threats, alert activists and researchers to these, and make our resource pool widely available to collaborators.”¹⁰⁸

Within this broad framework, organisations producing investigative reports are positioned differently in relation to the State. There is contrast between a specialist news publisher like [Bellingcat](https://www.bellingcat.com/)¹⁰⁹ which focuses on publicly available data and open-source investigation as an entry point for establishing facts, and investigative entities that work closely with particular state agendas and/or are aligned with particular companies.

4.2.3 What output do investigative responses publish?

Investigative reports take a range of forms. Most NGO and journalistic investigations focus on providing information on the mechanics and sources of disinformation (to expose approaches to creating and distributing disinformation) and creating in-depth reporting for broad consumption. Transparency on methods (and indeed co-creation and participation via crowd sourcing in evidence-gathering) is a key part of OSINT (Open Source Intelligence) approaches, as practiced by BellingCat and other hybrid organisations, while some entities produce data based on their specialities (for example, social graph network analysis methods in the case of Graphika). Investigations by internal groups within internet communications companies do not typically provide complete data on how they identified disinformation. It is an ongoing critique of the companies’ approaches to their own identification of disinformation, as well as their support to others, that there is a lack of provision of data to help identify, categorise and define disinformation campaigns.

However, a growing number of groups are trying to establish shared methodologies for classification. One example is the work of the Digital Forensic Research Lab to develop a framework for categorisation in their Dichotomies of Disinformation project (Brooking, et al., 2020) (with support from Jigsaw, a division of Alphabet, the holding company of Google). This categorisation approach includes over 150 binary, text-based and quantitative variables grouped under a top-line set of variables that includes: target, platform, content, method, attribution and intent. Other categorisation work includes the Computational Propaganda Project’s surveys of organised social media manipulation based on an assessment of news media reporting (Bradshaw & Howard, 2019), and the Institute for the Future’s reporting on types of state-sponsored trolling within disinformation and online harassment campaigns (Monaco & Nyst, 2018).

¹⁰⁸ <https://www.disinfo.eu/>

¹⁰⁹ <https://www.bellingcat.com/>

4.2.4 Who are the primary actors and who funds investigative responses?

A range of initiatives work on organised investigations into disinformation and produce in-depth reporting. These include:

Entities with a primary focus on disinformation: The Digital Forensic Research Lab of the Atlantic Council is an example of an organisation with a strong focus on identifying, unpacking and countering disinformation campaigns. It publishes reports, develops field expertise and identifies methodologies and tracking approaches (Brooking, et al., 2020). In the European context, EU Disinfo Lab is a more recently established NGO designed to maintain both a platform for analysis of disinformation in Europe, although it also monitors and identifies disinformation operations with international dimensions (EU Disinfo Lab, 2019a; Carmichael & Hussain, 2019). Actors responding to disinformation in this subcategory span foundation and government-funded outfits, non-governmental organisations, and dedicated academic programmes. Some investigations have delved into the business models used by many internet companies, attributing to these a propensity towards the propagation of rumour and conspiracy theorists. For example, the Center for Humane Technology says that YouTube recommended conspiracy videos by Alex Jones more than 15 billion times.¹¹⁰

Entities with methodologies relevant to disinformation, such as Open-Source Intelligence (OSINT): In parallel with the development of the disinformation research and investigation field in the past six years, there has also been the growth of an increasingly robust field of open-source investigation more broadly, using 'open source' and social media sources to conduct investigations into topics such as war crimes and chemical weapons usage. An example of an organisation in this field is [Bellingcat](#), self-described as an "...independent international collective of researchers, investigators and citizen journalists using open source and social media investigation to probe a variety of subjects – from (...) drug lords and crimes against humanity, to tracking the use of chemical weapons and conflicts worldwide."¹¹¹ It has staff and contributors in more than 20 countries around the world, who work at the intersection of advanced technology, forensic research, journalism, investigations, transparency and accountability monitoring. Entities in this group include foundation and government-funded outfits, and NGOs.

Investigations by existing non-governmental watchdogs or monitors with a thematic or sectoral Freedom of Expression focus: Although disinformation should not be conflated with hate speech, the combination of the two involves a range of existing groups who investigate patterns of malicious information-sharing in particular thematic contexts. One example is the [Southern Poverty Law Center](#) in the U.S., which exists to "monitor hate groups and other extremists throughout the United States and expose their activities to the public, the media and law enforcement."¹¹² They provide a comprehensive biannual report into the status of these movements and their activities, as well as specific reports into particular propaganda activities. Similarly, the London-based Institute for Strategic Dialogue (ISD)¹¹³ documents and produces reports on extremist violence and related speech. (As chapter 7.1 outlines, there are significant overlaps between normative and ethical responses to disinformation, and the issue of hate speech).

¹¹⁰ <http://humanetech.com/wp-content/uploads/2019/07/CHT-Undivided-Attention-Podcast-Ep.4-Down-the-Rabbit-Hole.pdf> ; <https://www.newamerica.org/oti/events/online-getting-to-the-source-of-the-2020-infodemic-its-the-business-model/>

¹¹¹ <https://www.bellingcat.com/about/>

¹¹² <https://www.splcenter.org/fighting-hate>

¹¹³ <https://www.isdglobal.org/isdapproach/>

In-depth investigations by news outlets: A range of news outlets maintain ongoing disinformation investigatory beats. One example is BuzzFeed News, providing insights and investigations into individual disinformation campaigns and trends in disinformation, such as the growing use of pay-for-hire PR firms in disinformation (Silverman, et al., 2020). Other outlets have conducted in-depth investigations of particular campaigns, such as Rappler’s mapping of disinformation networks during and after the 2016 presidential elections in the Philippines (Ressa 2016; Posetti et al., 2019a), the work of South Africa’s Daily Maverick referenced above (Thamm 2019; Posetti et al., 2019a), and that produced by the African Network of Centers for Investigative Reporting on media manipulation in South Africa¹¹⁴ (ANCIR, n/d). Another contribution comes from CodaStory, which has a disinformation specialisation and focuses on investigations into orchestrated campaigns connected to state actors and disinformation agents for hire (Dorroh 2020).¹¹⁵

Action-oriented academic research: A burgeoning number of academic departments produce both meta-analyses of disinformation campaign strategies, for example the inventories of organised social media manipulation (based on news media content analysis) from the Computational Propaganda Project at the Oxford Internet Institute (Bradshaw & Howard, 2019), as well as detailed research into specific strategies and country-contexts. An example of the latter is academic work on networked social media manipulation in the Philippines (Ong & Cabañes, 2018). Other research from the Reuters Institute for the Study of Journalism at the University of Oxford focuses on the exposure of the public to disinformation in a number of countries, as well as investigative responses of journalism (Posetti et al., 2019a) and media effects (Nielsen & Graves, 2017).

Commercial entities working in social network analysis and cyber-security: A range of commercial companies provide services or conduct investigative research into disinformation campaigns. An example is Graphika, which focuses on detecting “strategic influence campaigns online and at scale by analyzing network anomalies and identifying objects propagating through network maps with a high degree of social contagion that are likely to quickly reach virality.”¹¹⁶ The company applies social media network analysis to conduct investigations into specific campaigns. These investigations can be in coordination with other actors - for example, with companies such as Facebook, in the ‘Operation #FFS: Fake Face Swarm’ (Nimmo et al., 2019b), an analysis of fake profiles/accounts tied to the Epoch Media group. Another example of a commercial entity is FireEye, which has a commercial cybersecurity background. It has identified and investigated cybersecurity breaches and related disinformation campaigns originating in various States (Revelli & Foster, 2020).

Investigations by internal company threat mitigation teams: All major social media companies have internal threat analysis teams, and teams dedicated to ‘site integrity’ or identifying ‘coordinated inauthentic behavior’ (Gleicher, 2018a). For example, Facebook has produced a report on tackling co-ordinated inauthentic behaviour in a number of countries (Gleicher, 2020). These teams sometimes share specific data to outside partners or collaborate/contract with external companies and non-profit/academic groups (Gleicher, 2018b). In the case of Facebook this includes collaborations with a number

¹¹⁴ <https://s3-eu-west-1.amazonaws.com/s3.sourceafrica.net/documents/118115/Manufacturing-Divides.pdf>

¹¹⁵ See also this video panel discussion about in depth journalistic investigations into disinformation in the context of COVID-19 featuring CodaStory Editor Natalia Anteleva, BuzzFeed’s disinformation specialist Jane Lytvynenko, and Rappler’s Executive Editor, Maria Ressa: https://www.youtube.com/watch?v=tBp4OKSW_ho&feature=youtu.be

¹¹⁶ <https://www.graphika.com/graphika-labs>

of the other types of entities cited in this chapter - e.g. Digital Forensic Research Lab, Graphika, FireEye.

As can be seen from these examples, investigative reporting on disinformation is funded by a range of actors. Non-profit and non-governmental actors receive a combination of foundation funding, corporate and state funding. Some actors are more institutionally positioned in this respect - for example the Digital Forensic Research Lab at the Atlantic Council is part of a larger entity that receives significant funding from the British, U.S. and UAE governments, and additional backing from Facebook (Lapowsky, 2018). Other entities like Bellingcat are funded by foundations and provide training support and workshops to supplement their core income. Some legacy news organisations are also involved in collaborative investigative work on disinformation which attracts donor funding (e.g. through the International Consortium of Investigative Journalists), while others undertake independent investigations consistent with a mission for journalism designed to hold power to account.

Collaborative and interdisciplinary investigative responses – for example combining the expertise of actors in several of the categories above can heighten the effectiveness of these interventions. For example, Rappler’s journalistic investigations in the Philippines have involved partnerships with NGOs, academics and technology experts.

4.2.5 Response case study: COVID-19 disinformation

Due to their more in-depth and resource-intensive nature, and the short timeline of the pandemic, by May 2020 there were fewer published investigative responses to COVID-19 compared to more straight-forward fact-checking and verification efforts. Nevertheless, organisations specialising in investigative responses included outputs from several NGOs¹¹⁷, news publishers (Evans, 2020), think tanks (EUvsDisinfo, 2020), and joint investigations between academics and independent media (Hollowood & Mostrous, 2020). Topics being investigated include COVID-19 disinformation campaigns launched by state-sponsored media, violent extremist movements, anti-migrant, and far-right networks (ISD, 2020a). These operate across key social communications companies, including Twitter (open posts and direct messaging), Facebook (including profiles, groups, pages, Messenger), YouTube (videos and comments), WhatsApp, and Instagram (open posts and private messaging), despite efforts of these companies to counter the ‘disinfodemic’.

Most analysis to date does not involve in-depth investigation by foundations, think tanks or commercial entities, but reporting by news outlets, for example from ProPublica (Kao, 2020) and the *New York Times* (New York Times, 2020) in the U.S., and Rappler (Gaw, 2020) in the Philippines. The overt and continuous spread of disinformation by political leaders during the pandemic has been extensively reported in the media, along with assessments of how statistics are instrumentalised and used to convey misleading impressions.

Another category of investigative responses to COVID-19 disinformation includes guidance on types of disinformation identified to date, such as two policy briefs about the ‘disinfodemic’ published by UNESCO in partnership with the International Center for Journalists (ICFJ) (Posetti & Bontcheva, 2020). These identified nine types of COVID-19 era disinformation, four main vectors, and ten modalities of response. See also research from

¹¹⁷ <https://rsf.org/en/disinformation>

the Reuters Institute on COVID-19 disinformation types, sources and claims (Brennen et al., 2020) which identified political leaders and celebrities as top sources of disinformation.

Within internet communications companies, internal threat mitigation teams, either working independently, or in tandem with other expert actors, were also undertaking investigations into COVID-19 disinformation (Shu & Shieber, 2020). The results disclosed have been piecemeal,¹¹⁸ and specialist journalists have found them wanting (Turvill, 2020).

4.2.6 How are investigative responses evaluated?

Many actors are transparent on methods and processes and publish publicly accessible reports on their findings. However, explicit evaluations of impact and effectiveness are not publicly available from most of the actors involved in investigative reporting on disinformation. One area of visible results is in the context of industry-driven and collaborative investigations of disinformation campaigns - where specific takedowns of accounts and content related to an investigation occurs on Facebook, Twitter or another social media platform. Similarly, in the context of government-commissioned work, for example into foreign interference in the U.S. 2016 elections, data is directly fed into Congressional hearings.

4.2.7 Challenges and opportunities

Investigative reporting moves beyond individual fact-checking and debunks to produce deeper insights and analysis as well as details on specific campaigns. As this field has matured there is a growing ability to track disinformation actors over time. See, for example, the ongoing tracking of innovations or approaches in reported foreign interference in the U.S. elections 2016 U.S. elections (U.S. Senate Select Committee on Intelligence, 2018) through to campaigns such as IRA CopyPasta (François, et al., 2019).

A challenge to note is that journalists conducting investigations into disinformation are vulnerable to attacks against them, such as online harassment and targeted disinformation about them, as in the case of Maria Ressa at Rappler (Posetti, 2017). A number of internet communication companies have offered a degree of support such as the Facebook - Committee to Protect Journalists safety tips to protect sources and contacts¹¹⁹, and Google's Project Shield¹²⁰. However, there has been criticism of tardy company responses to complaints of harassment, and to making it the responsibility of the victim to protect themselves by blocking, reporting and deleting rather than the company taking swift action (Posetti, 2020).

As organisations move to codify and quantify the nature of disinformation campaigns, a body of data is developing that enables comparative analysis (as noted above). More organisations also engage in public education alongside intensive report-writing and investigations in order to ensure sharing of good practices and new approaches to countering disinformation. Examples of this include Digital Forensic Research Lab's annual

¹¹⁸ https://en.unesco.org/sites/default/files/unesco_covid_brief_en.pdf

¹¹⁹ <https://www.facebookblueprint.com/student/path/188883-journalist-safety>

¹²⁰ <https://blog.google/outreach-initiatives/google-news-initiative/our-efforts-help-protect-journalists-online/>

Digital Sherlocks methods-sharing conference¹²¹, EU Disinfo Lab's annual conference, as well as Bellingcat's open-source methods and training.

However, in-depth investigations face significant challenges beyond cost and complexity. Most investigations are conducted without access to the complete data sets necessary to fully understand a particular campaign as internet communications companies do not routinely provide this data. Twitter has explained its data disclosure policy in an article by its Head of Site Integrity (Roth, 2019), and Facebook has been criticised by researchers for delays in providing data access but has recently released a larger data set in line with its commitments (King & Persily, 2020). Another issue is the restriction of access to a limited number of researchers, who are also frequently the recipients of large grants from these companies.

Researchers also have limited information and tools to do cross-platform analysis, despite the fact that few organised disinformation (or viral misinformation) efforts are restricted to one single platform. A particular problem is the issue of accessing data on information shared on messaging apps - where disinformation is known to proliferate - which are often end-to-end encrypted for reasons of security and privacy. However, these companies do have access to metadata on traffic and groups, even if they do not have access to specific messages. Access to this information could help investigators to detect patterns of activity by disinformation networks.

In the past four years there has been a heavy initial focus on disinformation deemed to be sponsored by one State in particular. However, as outlined above, recent corporate, academic and investigatory responses are starting to focus on a wider range of States and private/governmental actors involved.

Similarly, as noted above, there are significant gaps in access to information to adequately support civil society, journalism and academia to understand cross-platform as well as messaging-based disinformation campaigns. This represents an opportunity for internet communications companies to collaborate with researchers and civil society organisations with specialist skills in this area on data analysis and policy development.

4.2.8 Recommendations for investigative responses

A number of recommendations for action can be adduced from this chapter for a range of actors. They include:

- All stakeholders could recognise the need to invest in critical, independent investigative journalism as a defensive measure against disinformation, particularly as COVID-19 financial pressures deliver death blows to news outlets around the world and threaten costly investigative journalism initiatives.
- Internet communications companies could provide broader and better access to their datasets to independent researchers studying disinformation, including those who do not receive significant research funding from these companies, in the interests of knowledge sharing to combat disinformation

¹²¹ <https://digitalsherlocks.org>

Donors and research organisations could:

- Increase investment in interdisciplinary and collaborative investigations, fostering cooperation between academic researchers, commercial data scientists, NGOs and news organisations.
- Fund quick-turnaround disinformation investigations during emergency situations such as the COVID-19 crisis