A Contextual Inquiry of The International Fact-Checking Network and Factuality on Social Media

Abstract

The post-truth era is co-located with the rise of social media. After years of avoiding action on posts based on accuracy of content, in 2016, Facebook announced that it would partner with fact checkers who were signatories of the International Fact-Checking Network's (IFCN) code of principles. This partnership was an admittance to a regulatory ideal of non-subjectivity, for at least a selection of posts. Other platforms followed suit in partnering with 'third party fact checkers' certified by IFCN, making IFCN's guidelines an indispensable infrastructure for content moderation on social media. Interrogating this infrastructure then enables us to trace the emerging boundary of a 'fact' on social media. We do this by tracing the geographical and political context in which IFCN emerged and contrast it with the social, political and material realities of social media users outside that context.

In particular, we share findings from annotating claims in social media posts from one such context in India. Reflecting on our process of annotating 2200 multimodal posts from an Indian social media platform, we ask- what is the modality of a claim in multimedia user-generated content? We aim to show that the aesthetic and modality of content bears on what is considered to be a claim and ultimately a question of fact. This opens avenues to reclaim media artefacts from the terrain of subjectivity, into the boundaries of factuality.

'Post-Truth' in the Mainstream

In November 2016, Oxford Dictionaries named 'Post-Truth' the Word of the Year (OUP 2016). The term, an adjective, was defined as 'relating to or denoting circumstances in which objective facts are less influential in shaping public opinion than appeals to emotion and personal belief'*(ibid)*. As an example of its usage, the dictionary states "'in this era of post-truth politics, it's easy to cherry-pick data and come to whatever conclusion you desire.'"[1] The announcement tracked the term to a 1992 essay. The philosophical and political conceptualizations of the term however run much deeper.

[1]https://www.lexico.com/definition/post-truth

The Social Studies of Science (STS), as a discipline, has richly debated the social construction of 'truth' and 'facts'(Sismondo 2017). Similarly, critical analysis of journalism has contended with the notion of neutral or objective reporting of facts (Maras 2013). These debates raged in academic groups for many decades. Yet, as the Oxford announcement noted, in 2016 "Post-truth (went) from being a peripheral term to being a mainstay in political commentary" (OUP 2016).

Post-truth increasingly became a moniker to describe Donald Trump's campaigning success. The era of post-truth, however, is fundamentally linked to the evolution of the media into "social" media (Sawyer 2018). *Facts* emerge from particular configurations of practices, discourses, epistemic politics and institutions (Sismondo 2017). The list of constantly updating social media posts on a Facebook users home page, called a 'newsfeed'[2] presented an alternative configuration to that developed by legacy media organizations and knowledge institutions. In the social media age, factual information is arrived at by a "network of agreement" that span the entire globe, creating a "transnational space" (Sawyer 2018) that was instrumental in catalysing large-scale political change.

Introduction of Fact-Checking to Online Platforms

Platforms had long evaded any intervention on posts based on accuracy or veracity. Through 2016, Facebook's Community Standards allowed reporting and take-downs only if a post impinged on user safety through sexual exploitation, hate speech and bullying, or any other activities that presented a direct threat to personal safety (Facebook 2016). By the end of 2016, however, the platform changed its position. Almost a month after Oxford Dictionaries' announcement, Facebook announced that it would partner with fact checking organizations that had signed on to the International Fact-Checking Network's (IFCN) Code of Principles (Silverman 2016).

The Poynter Institute in the United States launched the International Fact-Checking Network in 2015 [3]. IFCN is cen-

[2]https://www.facebook.com/help/1155510281178725/how-news-feed-works

[3]https://www.poynter.org/ifcn/

tered around a 'code of principles' "which are a series of commitments organizations abide by to promote excellence in fact-checking".[4] Groups from across the world apply to be signatories of this code. The IFCN reviews applicants for compliance with the code and verifies signatories who pass the audit.

Facebook clarified that the partnership would only target intentional hoaxes while excluding claims from politicians or partisan disputes. "The worst of the worst" posts were eligible to be adjudicated against facts (Silverman 2016). This partnership was thus an admittance to a regulatory ideal of non-subjectivity, for at least a selection of posts.

At the time of announcement in 2016, Facebook did not pay fact checking organizations for their participation but expected the additional visibility would benefit these organizations. In 2017, Facebook volunteered funding. Some fact checking groups rejected the funding for concerns of journalistic independence (Ananny 2018).

Google followed Facebook and announced a partnership with IFCN with the goal that "content on Google Search and Google News (be) accurately fact checked" (Anderson 2017). In subsequent years, the ambit of fact checking would expand from the "worst of the worst" posts to include 'misleading' and 'out of context' posts shared by social media users. In 2019, when asked in Congress as to why "the Daily Caller, a publication with well-documented ties to white supremacists," was an official fact-checker for Facebook, Mr Zuckerberg deferred to the selection process of the IFCN, and emphasized that the body was independent of Facebook (Klepper 2019).

IFCN Guidelines as Fact Making

Between 2016 and 2021, ninety two organizations across the world would apply to be signatories to IFCN[5]. The certification became a necessary qualification for fact-checking organisations to partner with social media platforms. Newer platforms have inherited practices from their older counterparts and continue to rely on the IFCN certification for assessing fact checking expertise. While the exact terminology might vary across groups, roughly speaking, fact-checking groups mark posts on a spectrum from 'false' to 'true', using terms such as 'misleading' or 'partly true' for content in between those extremes. Some groups will additionally mark claims that cannot be verified as 'unverified'. [6] In admitting and relying on IFCN certified fact checkers, platforms conceded some terrain to fact making by expertise. Users' speech, free as it may be, became subject to some verification and penalties.

IFCN principles are now firmly embedded in platform content decisions. They present a counter force to "the archive of facts...developed by networks of (user) agreement" (Sawyer 2018). This counter-force, however, comes with its own practices, discourse, institutional and political context. Interrogating these practices and context allows us to understand the emerging boundaries of factual claims on

[4]https://www.ifcncodeofprinciples.poynter.org/
[5]https://www.ifcncodeofprinciples.poynter.org/signatories
[6]https://newsmeter.in/fact-check-methodology/

social media that are exempt from users' subjective interpretations. The pervasiveness of IFCN principles makes it especially important to do so. We do so by locating IFCN in its geographical, political and institutional context, but by first tracing the disciplinary boundaries of 'fact-checking'.

The Aims of Fact-Checking

Generally understood as an exercise of debunking, fact-checking traces the spread of political statements, reports and public domain knowledge by following the trail of publicly made claims across the various "media-political landscapes" that they traverse, in order to discover their origins. Methodologically, fact-checkers often rely on "factual coherence" to ascertain the veracity and truth value of a public claim (Graves 2017).

Fact, as an ontological category, has been positioned as fundamentally opposed to values and distinct from theoretical and speculatory knowledge claims. Philosophically speaking, facts are entities that are grounded in reality or in a "state of affairs." Facts are what correct judgements and make propositions true (Mulligan and Correia 2007). Simply put, facts correspond to truth; "there can be no difference in truth without a difference in facts"*()*. A belief is only considered true if it corresponds to a fact, and not otherwise (David 2020).

As per (Graves 2017), fact-checking work seems to adhere to scientific principles for producing verifiable, replicable and transparent knowledge. Truth-making within contemporary journalism is fundamentally scientific then, both in its practices and operative values. However, sociological critiques of science have pushed for the interpretive properties of scientific facts, demonstrating their socially constructed nature and the subsequent possibility of multiple coexisting interpretations (Pinch and Bijker 1984). Truth is a socially constructed entity, emerging from specific contexts and as a result of diverse meaning-making processes which shape social reality. (Shapin 1994) offered a social account of truth-making which insisted on approaching these questions with the lens of collective agency, participation and group judgements.

(Godler and Reich 2013) posit that in gathering and reporting news, journalists act pragmatically — they acknowledge the incompleteness of truth but nevertheless considers it a valuable goal to strive towards. While it may appear that journalists lean towards being social constructivists, it might be more useful to understand their practice as following a coherentist conception of knowledge. In other words, journalists operate on beliefs, using which they validate new information; the practice of objectivity necessitates interpretive flexibility for journalistic practice *(ibid)*. Therefore, objectivity is not always demonstrable in the process of seeking truth, but is considered a regulatory ideal for the practice of journalism.

Some, such as (Uscinski and Butler 2013), remain critical of the epistemic foundations of fact-checking, finding the distinctions between 'true' and 'false' to be over-simplified. (Graves 2017), however, argues that fact-checking strengthens the case against the "naive empiricism" of assuming a direct correspondence between news and truth/reality. Fol-

lowing an ethos of journalistic objectivity, fact-checkers use interpretative and triangulation methods to navigate a divided political terrain. Fact-checking work strives towards distilling the political from within the factual, underlining the socially constructed nature of truth-making as well as the political nature of facts.

IFCN- A History

To trace the history of IFCN, we rely on Poynter Institute's narration. "Fact-checking in its current incarnation was kick-started in the United States with the launch of Factcheck.org in 2003" (Mantzarlis 2016). In 2008, Politifact won the Pulitzer for its reporting of the 2008 US election (Adair 2009). During the 2008 campaigning, Politifact 'fact checked' 750 political claims. Subsequently, Politifact began fact checking statements made by congresspersons, the White House and television news hosts *(ibid)*.

A 2016 article by the Poynter Institute notes that fact-checking "is very much a digital movement" (Mantzarlis 2016)- Politifact owed its speed and scale of fact checking to using the "power of the World Wide Web" effectively *(ibid)*. The article also emphasized that despite its geography of origin, the field was becoming far less "Americanocentric". Yet, the fact-checking methodology and consequently the code of principles were designed and refined over a specific domain of content with certain defining features. One feature was the modality of the claims- IFCN style fact checking emerged from checking verbal and textual claims. The second feature was the salience of the people making the claims. Politifact was primarily directed towards checking claims made by people in the public eye.

Facts in User Generated Content, Multi-Modal Content

We juxtapose the aforementioned history with annotations of factual claims in 2200 user generated Hindi multi-media posts (accepted as dataset paper in ICWSM 2021). The posts were sampled from a bigger dataset containing posts categorized under 'Health' and 'Politics' on an Indian social media platform called ShareChat. The data spans a nine month period from March - December 2020. The platform has over 150 million monthly active users and operates in several regional Indian languages. The platform allows users to create and add filters to audio-visual content and share it with a text descriptor. More than half the posts contained images, a fourth contained videos and less than a tenth were only text posts. Since images and videos were always shared with tags or text descriptors, the posts in this dataset were multi-modal.

The annotation process borrowed methods from grounded theory to explore and make sense of the dataset. After familiarising ourselves with the data, we used 'open coding' to break down and analyse it, drawing out "embedded phenomena, patterns, concepts, and themes" (Matthew and Price 2010) that stood out to us. Next, using 'axial coding', we refined the broad groupings we developed in the 'open coding' phase under the five broad categories:

1. Implied/explicit source of the content

2. Type of Factual Claim

3. Intentionality Portrayed in the Video, if the post contains a video

4. Presence of Non Manipulated Images

5. Political Memes

The procedure of annotation has been described in detail in forthcoming work *(ICWSM 2021)*. This paper is concerned with the overlap between two categories of annotation and its implications for fact-checking work. 6% of posts that contained a factual claim were also labeled as a meme. 28% of posts that were labeled as memes (category 5) also contained a factual claim (category 2).

Digitally Circulated Memes- An Overview

Memes can be several different types and genres of things, from fashion trends to cuisines, from political opinions to architectural styles and so on. Given the broad range of ideas that the term seeks to encapsulate, the definition of memes is often vague and loosely articulated. The memetic object is a nebulous entity. (Shifman 2013) defines digitally circulated memes as "units of popular culture that are circulated, imitated, and transformed by individual internet users, creating a shared cultural experience in the process." The replication in Internet memes, (Shifman 2013) argues, can happen on three axes, namely content ("a specific text, referencing to both the ideas and the ideologies conveyed by it"), form ("physical incarnation of the message ... which is perceived through senses") and stance ("information memes convey about their own communication"). Stance captures the relationality that memes allow to emerge between individual expression and "vernacular creativity" (Milner 2013). It is through this dimension that memes come to embody and channel user agency, becoming vehicles and progenitors of genres of internet discourse. The element of stance lends to the narrative building capacity of memes.

What is by far the most definitive feature of memes is their significance for the formation of Internet communities and subsequently for political discourse. (Milner 2016) argues that memes are open and adaptable "objects of participation" that are produced collaboratively and bridge the gap between producers and consumers within emerging participatory digital cultures. Similarly, (Wiggins and Bowers 2015) consider memes as 'artifacts of participatory digital culture', contending that they are "messages transmitted by consumers–producers for discursive purposes" repeating subjects/themes that lie "within an established theme."

Memes are tools for expression and assertion, to negotiate existing socio-cultural norms (Gal, Shifman, and Kampf 2015) and enable participation in political discourse. This makes them socially constituted "performative acts" or "speech acts" imbued with intentionality (Grundlingh 2018). (Shifman 2013) emphasises the need to depict people as active agents for understanding Internet memes in particular, especially because meaning shifts dramatically during memetic diffusion online, but also because memes have significant political potential to trigger social change (Mick-Evans 2019). The dovetailing of user agency and political

expression is exemplified by the multilayered and shifting nature of meanings embedded within political memes.

With the propensity to be remixed, adapted, reworked and localised to diverse, but specific, socio-cultural contexts, memes can be understood as a shared or common language to communicate across geographies (Milner 2013; Knobel and Lankshear 2006). Memes reconstitute political communities and enable a creative engagement with public discourse. (Milner 2013) floats the idea of memes being "nationwide inside jokes", demonstrating their capacity to "balance the shared and the unique" and allowing "fixity and novelty" to be woven together with "cultural precedent and individual expression" through the memetic format. Political memes are multi-layered not only as cultural texts, but also as tools for building narratives and for political mobilisation.

Factual Claims in Memes

Memes capture information and meanings produced collaboratively through online interactions between users in networked discursive spaces and are relevant precisely because these interactions are an integral part of our collective digital heritage (Mick III 2019). As we see in our dataset, memes become the means to generate facts by a "network of agreement."

Figure 1 is an example of a post that was annotated as containing factual claims as well annotated as being memes. A number of claims are being made in this post- names are ascribed to photographs of people. Each of the four individuals have been identified with a political party. Certain actions and practices have been ascribed to the political parties. The three claims in unison provide background for the rhetorical question (listed at the bottom of the post) about the organizational ethos of two major political parties in the country.

बीजेपी पार्टी में सब हैं **वर्कर**,
कांग्रेस में एक परिवार के हैं सब **लेबर**

प्रियंका का विरोध   सोनिया का विरोध

विधायक अदिति सिंह निलंबित   विधायक राजेश प्रताप सिंह संस्पेंड

मोदी जी का विरोध

अरुण शौरी और यशवंत सिन्हा करते रहते है पर पार्टी के अभी भी है सदस्य

अब समझे आवाज कहां दबाई जाती है

Figure 1: Text Embedded in Image

The current methodological boundaries of fact checking would demand a flattening of this multi-modal post to a single dimension of claims that can be fact checked. The three claims would be treated as three separate textual claims and assessed independently. Any affective interaction between image and text as modality, or the three claims would be ignored.

Discussion

Platforms and media are slowly recognizing memes as a vector of misinformation (Snyder and Fischer 2021). Far from being the "worst of the worst" post that Facebook first aimed to target with its partnership with IFCN certified fact-checkers (Silverman 2016), a meme, in isolation, can appear inane. It's claim and effect comes from the broader participatory culture that it emerges from and is embedded in. We argue that it is within such a context that the memetic object should be examined and in relation to which its facticity can be understood as part of the broader discursive space in which it emerges.

Through this paper, we have attempted to emphasise that fact-checking arises from particular configurations of practices tied to a geographical and political context. Far from being self-evident, these practices emerged from a specific domain of checking verbal and textual claims made by prominent figures. Memes and other multi-modal content are force-fitted into these practices by compressing multiple media dimensions to a single one. But this does not tackle the overall claim behind such posts, defeating the broader goals of fact-checking.

With the rapid growth in TikTok like multi-modal content, it is important that fact checking methodologies adapt and expand to tackle multi-modal content such as image memes and short videos. A situated analysis of IFCN and dominant fact-checking practices on social media platforms shows that this evolution is possible.

Copyright

Acknowledgments

References

Adair, B. 2009. PolitiFact - PolitiFact wins Pulitzer.

Ananny, M. 2018. Checking in with the Facebook fact-checking partnership.

Anderson, E. 2017. Building trust online by partnering with the International Fact Checking Network.

David, M. 2020. The Correspondence Theory of Truth. In Zalta, E. N., ed., *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2020 edition.

Facebook. 2016. Community Standards.

Gal, N.; Shifman, L.; and Kampf, Z. 2015. "It Gets Better": Internet memes and the construction of collective identity. *New Media and Society*.

Godler, Y., and Reich, Z. 2013. How Journalists "Realize" Facts. *Journalism Practice* 7(6):674–689. Publisher: Routledge _eprint: https://doi.org/10.1080/17512786.2013.791067.

Graves, L. 2017. Anatomy of a Fact Check: Objective Practice and the Contested Epistemology of Fact Checking. *Communication, Culture & Critique* 10(3):518–537. ˍeprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/cccr.12163.

Grundlingh, L. 2018. Memes as speech acts. *Social Semiotics* 28(2):147–168. Publisher: Routledge ˍeprint: https://doi.org/10.1080/10350330.2017.1303020.

Klepper, D. 2019. Ocasio-Cortez questioned Zuckerberg about false political ads. Facebook is now defending its policy. Section: Politics.

Knobel, M., and Lankshear, C. 2006. Memes and affinities: Cultural replication and literacy education.

Mantzarlis, A. 2016. There's been an explosion of international fact-checkers, but they face big challenges.

Maras, S. 2013. *Objectivity in Journalism | Wiley*. Wiley.

Matthew, J., and Price, C. 2010. Gale eBooks - Document - Coding: Open Coding.

Mick-Evans, K. 2019. *One Does Not Simply Preserve Internet Memes: Preserving Internet Memes Via Participatory Community -Based Approaches*. Ph.D. Dissertation.

Milner, R. M. 2013. Media Lingua Franca: Fixity, Novelty, and Vernacular Creativity in Internet Memes. *AoIR Selected Papers of Internet Research*.

Milner, R. 2016. The World Made Meme | The MIT Press.

OUP. 2016. Oxford Word of the Year 2016 | Oxford Languages.

Pinch, T. J., and Bijker, W. E. 1984. The Social Construction of Facts and Artefacts: or How the Sociology of Science and the Sociology of Technology might Benefit Each Other. *Social Studies of Science* 14(3):399–441. Publisher: SAGE Publications Ltd.

Sawyer, M. E. 2018. Post-Truth, Social Media, and the "Real" as Phantasm. In Stenmark, M.; Fuller, S.; and Zackariasson, U., eds., *Relativism and Post-Truth in Contemporary Society: Possibilities and Challenges*. Cham: Springer International Publishing. 55–69.

Shapin, S. 1994. *A Social History of Truth*. University of Chicago Press.

Shifman, L. 2013. Memes in a Digital World: Reconciling with a Conceptual Troublemaker. *Journal of Computer-Mediated Communication* 18(3):362–377. ˍeprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/jcc4.12013.

Silverman, C. 2016. Facebook Is Turning To Fact-Checkers To Fight Fake News.

Sismondo, S. 2017. Post-truth? *Social Studies of Science* 47(1):3–6. Publisher: SAGE Publications Ltd.

Snyder, A., and Fischer, S. 2021. How memes became a major vehicle for misinformation.

Uscinski, J. E., and Butler, R. W. 2013. The Epistemology of Fact Checking. *Critical Review* 25(2):162–180. Publisher: Routledge ˍeprint: https://doi.org/10.1080/08913811.2013.843872.

Wiggins, B. E., and Bowers, G. B. 2015. Memes as genre: A structurational analysis of the memescape. *New Media & Society* 17(11):1886–1906. Publisher: SAGE Publications.