

## WORKING PAPER 32 – – **The Impact of Gonzalez v. Google on Safe Harbor for Intermediaries in India**

By Dushyant Kishan Kaul

11 September 2023- DOI: 10.48251/SADF.ISSN.2406-5617.WP32



Dushyant Kishan Kaul attended University of California, Berkeley; University of Cambridge; Jindal Global Law School. The author is an Advocate practicing civil-commercial and technology laws at the High Court of Delhi

### Abstract

*This essay analyses how the decision in Gonzalez v. Google, LLC<sup>[1]</sup> (Gonzalez) will affect the immunity from liability conferred upon intermediaries. It calls for increased platform responsibility through self-regulation and occasional state nudges, as opposed to curtailment of platforms' speech and expression through a content liability approach. Finally, it uses the latest Information Technology (IT) Amendment, 2021 in India as a frame to argue that a stronger stance in Gonzalez was needed considering other jurisdictions like India that are increasingly looking to hold platforms accountable for the content they host.*

Keywords: *algorithm, content, intermediary, immunity, notice-and-takedown*

### **Introduction**

The US, §230 of the Communications Decency Act, 1996 (CDA) provides immunity for any information provided by 'another' information service provider. It essentially provides a safe harbour to interactive computer services (ICS) by treating them as publishers of third-party information. This immunity is available under the title of 'Good Samaritan Blocking and Screening of Offensive Material'.

In *Gonzalez*, the Plaintiffs were the family members of Nohemi Gonzalez, a 23-year-old slain victim of the terrorist attacks perpetrated by the Islamic State of Iraq and Syria (ISIS) in Paris in November 2015. They argued that YouTube, despite being aware of the presence of videos propagating terrorism on their platform, allowed ISIS to exploit YouTube's algorithm to target and suggest those videos to interested users based on their viewing history. These algorithms are inherent to internet platforms and serve to sort content depending on relevance and user preferences. They accused Google, which owns YouTube, of being secondarily liable under §2333(a) and (d)(2) for 'aiding and abetting' terrorism by conspiring with ISIS. It was contended that YouTube was an essential forum for ISIS's brainwashing and recruiting people as well as a crucial space wherein terrorist attacks were planned. In this regard, the US Supreme Court heard an appeal from the decision of the Ninth Circuit, which had upheld the lower court's decision to bar the plaintiff's claim under §230.

The statute in question, the Communications Decency Act (CDA), 1996 is a pre-algorithm statute that could not possibly have contemplated having to deal with the question of whether targeted recommendations amount to publication. In particular, the present case involved the presentation and organization of thumbnails by an ostensibly content-neutral platform that actively selects the placement of videos on the website and recommends them to users through an operating algorithm that applies uniformly across all types of content. The question is whether affirmative acts by Interactive Computer Services (ICS) like Google, which don't just let this content stay on a website but actively recommend it, entitle them to immunity. In other words, it involves the role of Google in recommending ISIS videos based on users' search history and whether this amounts to assisting ISIS in an act of international terrorism.

The Ninth Circuit opined that there is no civil liability under the domestic anti-terrorism statute, namely the Anti-Terrorism Act (ATA). They held that there was no loss of immunity for making information available to the public or giving third parties a neutral

means to host content. Thus, even if ISIS videos were recommended to users, there was entitlement to immunity because this content was not treated any differently from other third-party content. Google's content-neutral algorithm was thus held to function as a traditional search engine in recommending videos based on users' search history.

### **The US Supreme Court in *Gonzalez*: A Missed Opportunity**

On May 18, 2023, the US Supreme Court, except on the plaintiff's claim of secondary liability for Google sharing advertising revenue with user accounts through which ISIS videos were uploaded, agreed with the Ninth Circuit, and held that most of the Plaintiff's appeal did not fall within the application of §230. It essentially declined to rule on the scope of §230 and remanded the matter *per curiam* to the Ninth Circuit to be considered in light of its decision in *Taamneh v. Twitter*.

The court would ideally have held that Google is entitled to immunity under § 230(c)(1) for the following reasons. First, the claim pressed for 'aiding and abetting' international terrorism under the Anti-Terrorism Act of 1990 (ATA) has a statutory requirement of proving knowledge, i.e., that the videos were purposely recommended. This presents a high threshold to meet as it is difficult to hold that neutral suggestions based on users' interests amount to 'aiding and abetting'. Such lawsuits portend a litany of complaints that may similarly seek to hold platforms liable and where the threshold for liability may not be as high. Second, the court could have held that immunizing targeted recommendations is not creating a new judicial standard but is simply interpreting § 230(c)(1) to include them within its ambit. This is because § 230(c)(1) should be purposively construed to protect a platform recommending videos through a catalogue of thumbnails. Third, recommendations using thumbnails should not be actionable as this is a central feature of how most platforms function. Holding platforms liable every time their algorithm directs

users to potentially defamatory or harmful content, alongside their inherently subjective nature, would hinder the smooth functioning of the internet. Fourth, declaring this system of sorting and prioritization as unlawful would interfere with platforms' choice of how content appears on their platform. In fact, it would open a Pandora's box wherein lawsuits would start questioning several inherent features and aspects of algorithms. This concern stems from the decision of the Fourth Circuit in *Henderson v. Source of Public Data*, where the court stripped the concerned entities of their immunity based on how third-party information was presented. It implicated companies for 'improper content' despite this flying in the face of § 230. This narrow reading of § 230 is insidious for it allows complainants to plead around platform immunity by focusing on how the content is put together rather than the source of the information (Goldman, 2022).

Moving forward, courts in the future will have to grapple with the interpretation of § 230 when claims emerge that do not involve the Anti-terrorism Act (ATA). The court needs to address the fact that the term 'neutral tool / algorithm' in content moderation is a misnomer – as all platforms discriminate in uploading content. The court should also retain § 230 for when authors pay for the content, as this is essential for the very existence of platforms (Goldman, 2023).

Allowing such appeals would result in the resurrection of decisions like in *Fair Housing Council of San Fernando Valley v. Roommates.com, LLC*. Roommates.com, where an automated web portal providing housing-related services, was accused of violating laws by having selection preferences based on users' race, religion, sexual orientation etc., to match compatible roommates looking for a place to stay. Holding Roommates.com liable for the 'creation' and 'development' of content, by making a 'material contribution' to illegal third-party content, adds uncertainty to what constitutes 'publication' under § 230. It discourages platforms from posting questionnaires intended to sort consumer preferences and improve user experience. The perverse incentive for platforms is that they can enjoy

immunity under § 230 by simply ignoring takedown requests. Surely, this type of behaviour ought to be discouraged.

Similarly, the grant of immunity under § 230 in *Herrick v. GRINDR* and its denial in *Lemmon v. Snap* involved the very same claim: product design defect. The first case involved malicious impersonation and catfishing by the victim's ex-boyfriend on a web-based dating application for gays and bisexuals; while the second case involved putting a speed filter on the multimedia instant messaging application 'Snapchat', and how that contributed to a car accident, and claimed three lives. In fact, the latter decision of the Ninth Circuit incentivizes claimants to plead around the immunity under § 230 by asserting negligence and tort-related claims. This may present a problem for courts in the future because the terms under which a product is considered defective to fall outside the scope of 'publication' under § 230, are unclear and vary from case to case. Undoubtedly, such occurrences indicate problems with trying to shoehorn § 230 into spaces that Congress did not envisage.

While Congress would face resistance if it decided to amend § 230 because of First Amendment rights and the consolidation of power among technology companies, the provision requires tinkering with. First, while § 230(c)(2) immunizes platforms for voluntary good faith policing efforts, providing them with immunity disincentivizes proactive content removal. As a result, platforms are encouraged to only fulfil the bare minimal policing whereas the reality is that in today's world, platforms should engage in extensive policing efforts. In fact, platforms may be incentivized to promote socially harmful content that may nonetheless be economically valuable to them. One possible solution is to impose a 'duty of care' standard upon platforms where immunity is claimed only by platforms undertaking reasonable measures to address content that it knows to be unlawful. Incorporating an element of reasonableness might lead to the 'reasonable steps' component being understood differently for big technology companies compared to start-

ups, which at their incipient stages enjoy limited funds and resources (Smith & Van Alstyne).

There are legitimate reasons for protecting internet platforms under the Communications Decency Act (CDA), including the need to promote free speech and prevent undue state censorship. While platforms cannot possibly review every post, if their algorithm targets recommendations in the same way for all types of content, this cannot attract liability. It would otherwise be impossible to draw the line whereby platforms may be held liable for amplifying any potentially controversial content. Platforms would then be sued simply for using ‘nudging’ tools. Other than that, there are problems of de-platformed users signing up under a different name and the fact that what amounts to terrorist-content changes depending on the context. Thus, denying immunity would detrimentally impact both the subject (ex: any form of controversial speech) and the object (search engines, small platforms, unpopular political groups, etc.).

To summarise, the court should nonetheless have addressed the following considerations: (i) the future implications for how platforms organize and present information, (ii) how § 230 should be interpreted to protect platforms from claimants seeking to plead around § 230 by pleading alternative claims of promissory estoppel, product design liability or negligence, and (iii) whether the term ‘publication’ provides a safety net for platforms who engage in promoting unlawful third-party content through methods like amplification, nudging, etc. Re-wording the statutory provision would promote the spirit of self-enforced content moderation whilst ensuring that platforms cannot avail of this safe harbour for their unlawful acts or omissions.

### **Impact of Gonzalez on the Notice-and-Takedown approach in India**

In India, platforms can avail of a safe harbour for hosting content under S.79 – subject to due diligence obligations introduced in the 2008 Amendment to the Information Technology (IT) Act, 2000.

The deleterious effect of skirting the issue of platform immunity in *Gonzalez* is that it fails to protect platforms in using algorithmic targeted recommendations in video placement and recommendations. The ramifications of this decision can be seen in the prospective implementation of the IT Rules, 2021,[2] in India. The Rules categorize platforms with over 5 million registered users as a ‘Significant Social Media Intermediary’ (SSMI) with additional obligations. Due diligence obligations apply upon receipt of actual knowledge pursuant to a court order, whereby content must be taken down within 36 hours. Moreover, the Rules also call for such intermediaries to identify the first originator of information for “serious offences”, which risks compromising the end-to-end encryption protection given by messaging platforms (Mehta, 2021).

Further, the Information Technology (IT) Rules, 2021 (R.3(1)(d)) have reinstated the requirement to remove content within 36 hours of having actual knowledge. This requirement is reduced to 24 hours when the matter concerns sexual imagery, nudity, or impersonation. The chilling effect is concerning because it would incentivize self-censorship to avoid liability. Moreover, Rule 6 empowers the government to designate any entity as an SSMI if a “material risk to harm” is demonstrated towards “state sovereignty, foreign relations or public order”. This excessive discretion makes it easy to justify decisions like banning TikTok due to the territorial dispute with China. While the government, in a press release, mentioned that Facebook, Instagram, Twitter, WhatsApp and YouTube would all fall within the category of SSIMs, many other user-generated content transmitters like LinkedIn, Medium, Reddit, TikTok and Twitch may also fall within this broad category. While large platforms like Twitter can easily comply, the compliance is burdensome if smaller platforms are designated as a ‘significant social media intermediary’ (SSMI).

Privacy concerns also arise when platforms must trace the first originator of information. For instance, WhatsApp would risk violating its own sacrosanct end-to-end encryption protection policy under Rule 4(2) whereby only the sender and recipient of information

would otherwise know the content of the message without any knowledge of the messaging service provider. While the method of tracing does not compel breaking end-to-end encryption, it is doubtful whether there is any other way of fulfilling this requirement. Identifying both the sender and recipient of the information necessarily requires tracing the information, which will lead to more control of the said information by platforms. Weakening data encryption will undermine users' data security, and platforms may risk violating their obligation to take reasonable data security measures. Dropping end-to-end encryption will only incentivize the handful of criminals among the millions of users on these platforms to use the dark web – or even create their own encrypted applications, like Al-Qaeda did in 2007. It is also unlikely that this provision meets the proportionality test of being a reasonable restriction to pursue the State's aim, which was instituted in *Puttaswamy v. Union of India*.

While voluntary self-moderation does not lead to a loss of safe harbour, concerns have arisen regarding how the government directly communicates with intermediaries – and has failed to provide originators with a hearing before imposing content restrictions. In fact, the methods of tracing first content originators are prone to easy circumvention and threaten to seriously undermine user privacy *en masse*. (Devadasan, 2023). Examples of potential State abuse are seen, for instance, when the State adopts “intimidation tactics” in raiding Twitter's office in Delhi after it tagged the posts of ruling leaders as “manipulated media” (Pfefferkom, 2021; Singh & Sekhri, 2021).

Such designated intermediaries are, under Rule 3(1)(b)(vi) of the Information Technology Rules, 2021, obligated not to host content that threatens State interests. There is little clarity regarding how this may be implemented, either by censoring the content altogether, or less restrictively, by flagging the content to disincentivize users from disseminating the information. To worsen the matter, the proposed rules establish a governmental fact-check unit to identify “fake or false or misleading information” with respect to any business. This might result in the creation of a censorial state which scrutinizes online content not just of

newspapers and social media platforms but also of service providers and file hosting companies. Neither “fake”, “false” nor “misleading” are defined – nor are the procedural safeguards for this unit set out. Given how content is governed based on the reasonable restrictions under Article 19(2) of the Constitution of India, factually inaccurate statements and other such content may not fit neatly within the categories set out thereunder. The *de facto* requirement of a state sanction vis-à-vis content through its fact-checking mechanism may itself result in the State indirectly dictating the kind of content that is shared on the platform. Arguably, this also raises concerns regarding the chilling effect on free speech u/A 19(1)(a). The government cannot possibly be the sole determinant of fake news and risks ending up as an Orwellian State (Vardarajan, 2023; Gupta, 2023).

### **Protecting Platform Immunity: The Way Forward**

While there should be greater platform accountability for its targeted recommendation algorithms, YouTube should not be stripped of its statutory immunity under § 230. Ranking and sorting algorithms are essential to the functioning of these websites so as to prioritize relevant content; these algorithms apply equally to all types of content. Imposing liability for amplification could possibly result in over-removal by platforms. Narrowing intermediary safe harbour by making platforms liable for their algorithmic recommender systems would detrimentally impact free speech (Giridhar and Devadasan, 2023).

Despite the robust culture in the US of protecting free speech under the First Amendment, the political and socio-economic leanings of regulators – coupled with the extent of influence exercised by the dispensation – may devastatingly stifle the flow of free speech. For instance, certain content may be controversial and may even be considered politically inflammatory. If the regulator is partisan or compromised, removal orders could lead to the internet being dominated by voices and filter bubbles from only one end of the spectrum.

Akriti Gaur innovatively postulates a framework of “induced self-regulation” which entails “regulatory nudges towards accountable self-regulation”. Platforms like Facebook and

WhatsApp have introduced fact-checking systems and bulk messaging limitations to counter fake news and other forms of misinformation. Exclusive platform self-regulation through advisory instruments[3] has seen little success, and government efforts face jurisdictional challenges as most big-tech giants are foreign entities. An environment of co-regulation fosters desirable government intervention without undermining users' free speech rights. To develop self-regulation norms, platforms should strengthen user autonomy, moderate content without impeding free speech and advocate for judicial oversight on state implementation through reasonable inspection mechanisms. An environment of collaboration is preferable to restrictive regimes that impose harsh penalties in using product designs and features (Gaur, 2021).

As Chinmayi Arun argues, a reformed platform regulation regime should displace the top-down approach of community standards and focus on bottom-up implementation. More sophisticated platform community rules and externalizing policymaking to third parties would better target local types of harmful content that often go undetected. Current third parties, however, focus only on regulating online speech. Bodies focusing exclusively on standard-setting thus become necessary. Ideally, co-development of guidelines by both the State and platforms should be encouraged by using third parties to provide independent perspectives that inform tailored and contextual platform responses. Liability for failing to act or engaging in censorial tendencies should, however, ultimately rest with the platform (Arun, 2018).

Monetization of user-generated content by platforms should thus not attract liability. Nonetheless, while platforms are incentivized to keep users engaged, a softer liability regime might be needed for websites actively promoting disinformation, inflammatory content, hate speech, harassment etc. A delicate balance must be achieved.

## **Conclusion**

Intermediaries are crucial in fostering communication through a peer-to-peer network among users. A ‘content liability’ model disenfranchises users from having more control over content dissemination. However, reduced platform liability without an increase in platform responsibility and transparency will fail to prevent platforms from employing processes that curb free speech as per private norms. There is a need for context-specific community guidelines that are based on the speech norms of local communities. To bridge the imbalance between platforms and users, enhancing user experience must be the priority in terms of the standards and practices adopted by platforms (Joshi, 2018). Depriving YouTube of its safe harbour under §230(c)(1) would be disastrous. But letting it completely off the hook is not desirable either.

## References

Arun, C. (2018, January 23). Rebalancing Regulation of Speech: Hyper-Local Content on Global Web-Based Platforms. Social Science Research Network papers. [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3108238](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3108238)

Bowers, J. & Zittrain, J. (2020, January 14). Answering impossible questions: Content governance in an age of disinformation. Misinformation Review. <https://misinforeview.hks.harvard.edu/article/content-governance-in-an-age-of-disinformation/>

Devadasan, V. (2023, January 18). Report on Intermediary Liability in India. Center for Communication Governance at National Law University, Delhi, <https://ccgnludelhi.wordpress.com/2023/01/18/report-on-intermediary-liability-in-india/>

Fair Housing Council of San Fernando Valley v. Roommates.com. Limited Liability Corporation 521 F.3d 1157 (9th Cir., 2008)

<https://law.justia.com/cases/federal/appellate-courts/ca9/04-56916/0456916-2011-02-25.html>

Gaur, A. (2021, March 2). Towards policy and regulatory approaches for combating misinformation in India. Information Society Project, Yale Law School. <https://law.yale.edu/isp/initiatives/wikimedia-initiative-intermediaries-and-information/wiii-blog/towards-policy-and-regulatory-approaches-combating-misinformation-india>

Giridhar, A. & Devadasan, V. (2023, February 23). YouTube lawsuit: Why the algorithm may not be the villain. Indian Express. <https://indianexpress.com/article/opinion/columns/google-youtube-us-supreme-court-lawsuit-8462436/>

Goldman, E. (2022, November 4). Fourth Circuit Takes a Wrecking Ball to Zeran and Section 230 – Henderson v. Public Data. Technology & Marketing Law Blog. <https://blog.ericgoldman.org/archives/2022/11/fourth-circuit-takes-a-wrecking-ball-to-zeran-and-section-230-henderson-v-public-data.htm>

Goldman, E. (2023, February 21). Quick Debrief on the Gonzalez v. Google Oral Arguments. Technology and Marketing Law Blog. <https://blog.ericgoldman.org/archives/2023/02/quick-debrief-on-the-gonzalez-v-google-oral-arguments.htm>

Gonzalez v. Google. US Supreme Court, No.21-1333 (2023). [https://www.supremecourt.gov/opinions/22pdf/21-1333\\_6j7a.pdf](https://www.supremecourt.gov/opinions/22pdf/21-1333_6j7a.pdf)

Gonzalez v. Google. Ninth Circuit, 2 F.4th 871 (2021). <https://law.justia.com/cases/federal/appellate-courts/ca9/18-16700/18-16700-2021-06-22.html>

Gupta, A. (2023, April 11). New IT Rules vest Centre with draconian powers of fact checking. The Indian Express, <https://indianexpress.com/article/opinion/columns/new-it-rules-centre-draconian-powers-fact-checking-8548308/>

Henderson v. Source of Public Data. 53 F.4th 110. (2022).  
<https://law.justia.com/cases/federal/appellate-courts/ca4/21-1678/21-1678-2022-11-03.html#:~:text=Here%20the%20district%20court%20erred,a%20publisher%20under%20the%20provision.>

Herrick v. GRINDR. 9<sup>th</sup> Cir., No.18-396 (2019). <https://casetext.com/case/herrick-v-grindr-llc-2>

Joshi, D. (2018, November 3). Beyond Intermediary Liability: Platform Responsibility for Harmful Speech in India. Observer Research Foundation. <https://www.orfonline.org/expert-speak/beyond-intermediary-liability-platform-responsibility-for-harmful-speech-in-india/>

Lemmon v. Snap. 9<sup>th</sup> Cir, No.20-55295. (2021).  
<https://law.justia.com/cases/federal/appellate-courts/ca9/20-55295/20-55295-2021-05-04.html>

Mehta, A. K. (2021, July 12). What about the Intermediary? Demystifying the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021. Supreme Court Cases (SCC) Online. <https://www.scconline.com/blog/post/2021/07/12/information-technology-intermediary-guidelines-and-digital-media-ethics-code-rules-2021-3/>

Taamneh v. Twitter. US Supreme Court, No.21-1496. (2023).

[https://www.supremecourt.gov/opinions/22pdf/21-1496\\_d18f.pdf](https://www.supremecourt.gov/opinions/22pdf/21-1496_d18f.pdf)

Pfefferkorn, R. (2021, March 3). New intermediary rules jeopardize the security of

Indian internet users. Brookings Tech Streams.

<https://www.brookings.edu/techstream/new-intermediary-rules-jeopardize-the-security-of-indian-internet-users/>

Puttaswamy v. Union of India; (2017) 10 Supreme Court Cases (SCC) 1

Singh, T. & and Sekhri, A. (2021, June 1). India's New Intermediary Guidelines:

The Gloves Come Off in the Assault on Civil Liberties. Verfassungsblog.

<https://verfassungsblog.de/it-rules-21/>

Smith & Van Alstyne. (2021, August 12). It's Time to Update Section 230. Harvard

Business Review Blog. <https://hbr.org/2021/08/its-time-to-update-section-230>

Vardarajan, A. (2023, April 7). Content Moderation and Speech Censorship – The

Latest IT Amendment Rules. Indian Constitutional Law and Philosophy.

<https://indconlawphil.wordpress.com/tag/intermediary-liability/>

---

[1] Docket for 21-1333 (Case Nos. 18-16700)

[2] Information Technology (Intermediary Guidelines and Digital Media Ethics Code), Rules, 2021, Ministry of Electronics and Information Technology, February 25, 2021

[3] For instance, the Information Trust Alliance (ITA) released a Code of Practice to develop a standard procedure on how to deal with complaints pertaining to fake news / content.

19 Avenue des Arts 2<sup>nd</sup> floor, 1210 Brussels, Belgium  
E 0833.606.320 RPM Bruxelles  
Email: [info@sadf.eu](mailto:info@sadf.eu) Web: [www.sadf.eu](http://www.sadf.eu)