

Senate Judiciary *Platform Accountability: Gonzalez and Reform*

Hany Farid, Ph.D.

Biography

Hany Farid is a Professor at the University of California, Berkeley with a joint appointment in Electrical Engineering & Computer Sciences and the School of Information. His research focuses on digital forensics, online harms, image analysis, and human perception. He received his undergraduate degree in Computer Science and Applied Mathematics from the University of Rochester in 1989, his M.S. in Computer Science from SUNY Albany, and his Ph.D. in Computer Science from the University of Pennsylvania in 1997. Following a two-year post-doctoral fellowship in Brain and Cognitive Sciences at MIT, he joined the faculty at Dartmouth College in 1999 where he remained until 2019. He is the recipient of an Alfred P. Sloan Fellowship, a John Simon Guggenheim Fellowship, and is a Fellow of the National Academy of Inventors.

Testimony

Background

In the summer of 2017, three Wisconsin teenagers were killed in a high-speed car crash. At the time of the crash, the boys were recording their speed of 123 mph on Snapchat's *Speed Filter*. This was not the first such incident. A 2015 crash left a Georgia man with permanent brain damage. Also in 2015 the *Speed Filter* was linked to the death of three young women in Philadelphia. And in 2017, five people in Florida died in a high-speed collision, which again reportedly involved the *Speed Filter*.

Following the 2017 tragedy, parents of the passengers sued Snapchat claiming that their product, which awarded “trophies, streaks, and social recognition,”

was negligently designed to encourage dangerous high-speed driving. In 2021, the Ninth Circuit ruled in favor of the parents and reversed a lower court's ruling that had previously emphasized the *Speed Filter* as creating third-party content and was thus deserving of Section 230 protection.

Section 230 immunizes platforms in that they cannot be treated as a publisher or speaker of third-party content. In this case, however, the Ninth Circuit found that plaintiff's claims did not seek to hold Snapchat liable for third-party content, but rather for a faulty product design which predictably encouraged dangerous behavior. In response, Snapchat removed the *Speed Filter*.

This landmark case – *Lemmon v. Snap* – made a critical distinction between a product's negligent design decisions and the underlying user-generated content.

Gonzalez

Frustratingly, over the past several years, most of the discussions of Section 230 – and most recently, in the US Supreme Court oral arguments in *Gonzalez v. Google* – this fundamental distinction between content and design has been overlooked and muddled.

At the heart of *Gonzalez* is whether Section 230 immunizes YouTube when they not only host third-party content, but make targeted recommendations of content. Google's attorneys argued that fundamental to organizing the world's information is the need to algorithmically sort and prioritize content. In this argument, however, they conflated a search feature with a recommendation feature. In the former, the algorithmic ordering of content is critical to the functioning of a Google search. In the latter, however, YouTube's *Watch Next* and *Recommended for You* features – which lie at the core of *Gonzalez* – are fundamental design decisions that materially contribute to the product's safety.

The core functionality of YouTube as a video-sharing site is to allow users to upload videos and allow other users to view and possibly search for videos. The basic functionality of recommending videos after a video is watched (*Watch Next*) and enumerating a list of recommended videos alongside each hosted video (*Recommended for You*) is a design decision made to increase user engagement and, in turn, ad revenue. While optimizing for these features may seem innocuous, this design decision has a critical safety flaw.

YouTube has argued that their recommendation algorithms are neutral in that they operate in the same way as it pertains to cat videos and ISIS videos. This is not the point. Because YouTube cannot distinguish between cat and ISIS videos, they have negligently designed their recommendation feature and should remove it until it operates accurately and safely.

YouTube has also argued that with 500 hours of video uploaded every minute, they must make decisions on how to organize this massive amount of content. But again, searching for a video based on a creator or topic is distinct from YouTube’s design of a recommendation feature where its sole purpose is to increase YouTube’s profits by encouraging users to binge watch more videos.

In doing so, the recommendation feature prioritizes the presentation of increasingly more bizarre and dangerous rabbit holes full of extremism, conspiracies¹, and dubious alternate COVID, climate-change, and political facts – content that YouTube has learned keeps users coming back for more.

Similar to Snapchat’s decision to create a *Speed Filter*, YouTube chose to create this recommendation feature which they knew, or should have known, was leading to harm.

By focusing on Section 230 immunity from user-generated content, we are overlooking product design decisions which predictably have – as is at issue in *Gonzalez* – allowed, and even encouraged, terrorist groups like ISIS to use YouTube to radicalize, recruit, and glorify global terror attacks.

Reform

During the recent *Gonzalez* oral arguments, Justice Kagan questioned if the Court or Congress should take up Section 230 reform, noting “we are not the nine greatest experts on the internet.” In this case, however, the issues do not require deep internet expertise (although I would advise the Justices (and Congress) to become more expert at all things technological).

The technology sector has been highly effective at muddying the waters and scaring the Courts and Congress with the claim that holding platforms responsible for what amounts to a faulty product design would destroy everything on the internet from a Google/Bing search to a Wikipedia page. This is simply untrue and we should not fall for this self-serving fear mongering.

While much of the debate around Section 230 has been highly partisan, it need not be. The core issue is not one of over- or under-moderation, but rather one of faulty and unsafe product design. As we routinely do in the offline world, we can insist that the technology in our pockets are safe. For example, we have done a good job of making sure that the battery powering our device doesn’t explode and kill or injure us, but have been negligent in ensuring that the software running on our device is safe.

The core tenants of Section 230 – limited liability for hosting user-generated content – can be protected while insisting, as in *Lemmon v. Snap*, that tech-

¹M. Faddoul, G. Chaslot, and H. Farid. A Longitudinal Analysis of YouTube’s Promotion of Conspiracy Videos. *arXiv*: 2003.03318, 2020.

nology that is now an inextricable part of our lives be designed in a way that is safe.

Summary

When, in 1996, Congress enacted Section 230 as part of the Communications Decency Act, they could not have envisioned today's internet and technology landscape. Congress could not have envisioned the phenomenal integration of our online and offline world, the trillion dollar sector founded on vacuuming up every morsel of our online (and offline) behaviors, and the highly-targeted algorithmic manipulation of nearly everything we see and consume online.

Nearly three decades later, we must rethink the overly broad interpretation of Section 230 that has moved from its original intent of not penalizing Good Samaritans to the current system of rewarding Bad Samaritans. This can be accomplished by clarifying that Section 230 is intended to protect platforms from liability based exclusively on their hosting of user-generated content, and not – as it has been expanded to include – a platform's design features that we now know is leading to significant harms to individuals, societies, and our very democracy.