

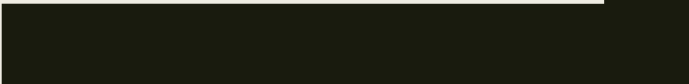


MODERATING DIGITAL PLATFORMS

US model: The first Amendment and Section 230 of
the Communications Decency Act

Beatriz García Quiroga

30 September, 2025



Today's Questions

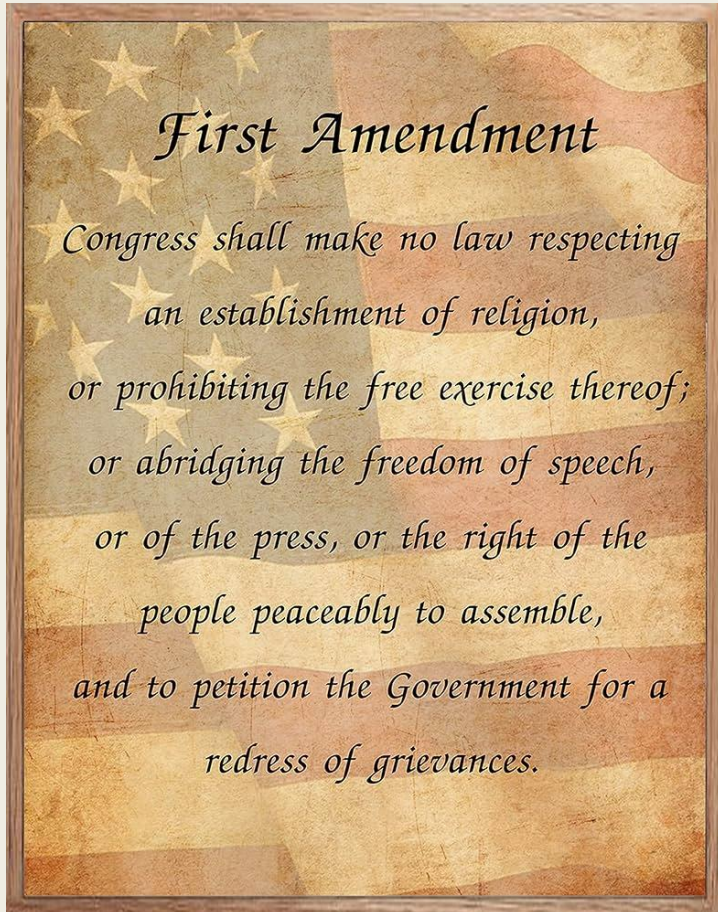
- What are the key US rules for online content moderation?
- Why were they developed?
- Are they still fit for today's digital world?

But before: overview of the U.S. Legal system:

- Constitution: Top legal source
- Federation: federal law > state law
- Common law: as laws are less developed, Courts have more power to interpret the law
- Politically appointed judges → ideology matters
- Regulation mainly ex post

¿WHAT ARE THE MAIN
RULES FOR CONTENT
MODERATION IN THE US?

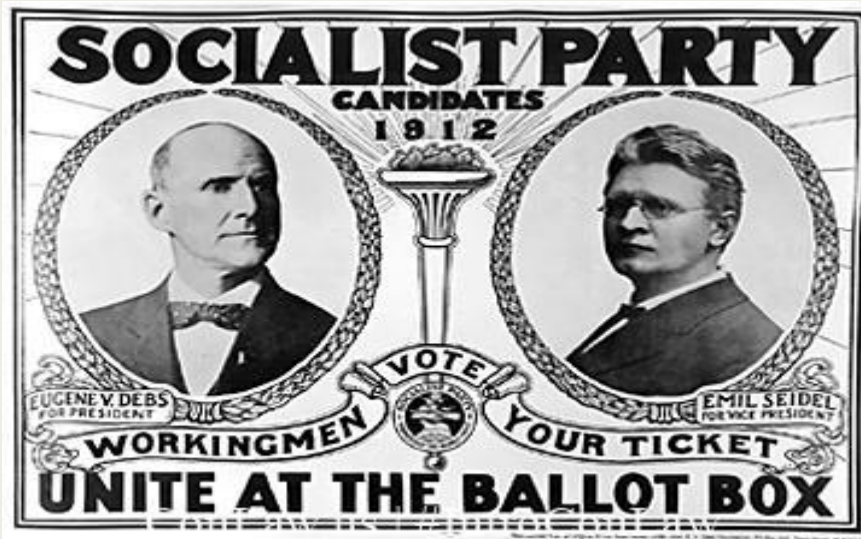
The 1st amendment (1791)



‘Congress shall make no law... abridging the freedom of speech, or of the press.’

The origin of current 1st Amendment doctrine

- Originally limited government censorship.
- Modern doctrine shaped after WWI.
- Goal: avoid the “chilling effect” on speech.



First pillar: Truth is better reached through the marketplace of ideas

The great Dissent Abrams V. US (1919)

“the ultimate good desired is better reached by free trade in ideas—that the best test of truth is the power of the thought to get itself accepted in the competition of the market, and that truth is the only ground upon which their wishes safely can be carried out”



Second pillar: lies about public figures are protected unless malicious

NY Times v. Sullivan (1964)

*new in the South, something understandable. . . .
Let Congress heed their rising voices,
for they will be heard."*

*—New York Times editorial
Saturday, March 19, 1960*

Heed Their Rising Voices

As the whole world knows by now, thousands of Southern Negro students are engaged in widespread non-violent demonstrations in positive affirmation of the right to live in human dignity as guaranteed by the U. S. Constitution and the Bill of Rights. In their efforts to uphold these guarantees, they are being met by an unprecedented wave of terror by those who would deny and negate that document which the whole world looks upon as setting the pattern for modern freedom. . . .

In Orangeburg, South Carolina, when 400 students peacefully sought to buy doughnuts and coffee at lunch counters in the business district, they were forcibly ejected, tear-gassed, soaked to the skin in freezing weather with fire hoses, arrested en masse and herded into an open barbed-wire stockade to stand for hours in the bitter cold.

In Montgomery, Alabama, after students sang "My Country, 'Tis of Thee" on the State Capitol steps, their leaders were expelled from school, and truckloads of police armed with shotguns and tear-gas ringed the Alabama State College Campus. When the entire student body protested to state authorities by refusing to re-register, their dining hall was padlocked in an attempt to starve them into submission.

protagonists of democracy. Their courage and amazing restraint have inspired millions and given a new dignity to the cause of freedom.

Small wonder that the Southern violators of the Constitution fear this new, non-violent brand of freedom fighter . . . even as they fear the upswelling right-to-vote movement. Small wonder that they are determined to destroy the one man who, more than any other, symbolizes the new spirit now sweeping the South—the Rev. Dr. Martin Luther King, Jr., world-famous leader of the Montgomery Bus Protest. For it is his doctrine of non-violence which has inspired and guided the students in their widening wave of sit-ins; and it is this same Dr. King who founded and is president of the Southern Christian Leadership Conference—the organization which is spearheading the surging right-to-vote movement. Under Dr. King's direction the Leadership Conference conducts Student Workshops and Seminars in the philosophy and technique of non-violent resistance.

Again and again the Southern violators have answered Dr. King's peaceful protests with intimidation and violence. They have bombed his home almost killing his wife and child. They have assaulted his person. They have arrested him seven times—for

of others—look for guidance and support, and thereby to intimidate *all* leaders who may rise in the South. Their strategy is to behead this affirmative movement, and thus to demoralize Negro Americans and weaken their will to struggle. The defense of Martin Luther King, spiritual leader of the student sit-in movement, clearly, therefore, is an integral part of the total struggle for freedom in the South.

Decent-minded Americans cannot help but applaud the creative *daring* of the students and the quiet heroism of Dr. King. But this is one of those moments in the stormy history of Freedom when men and women of good will must do more than applaud the rising-to-glory of others. The America whose good name hangs in the balance before a watchful world, the America whose heritage of Liberty these Southern Upholders of the Constitution are defending, is *our* America as well as theirs . . .

We must heed their rising voices—yes—but we must add our own.

We must extend ourselves above and beyond moral support and render the material help so urgently needed by those who are taking the risks, facing jail, and even death in a glorious re-affirmation of our Constitution and its Bill of Rights.

Third pillar: Lying is generally protected by freedom of speech – Congress cannot regulate against that



US v. Alvarez (2012)

"The remedy for speech that is false is speech that is true [...] Only a weak society needs government protection or intervention before it pursues its resolve to preserve the truth. Truth needs neither handcuffs nor a badge for its vindication."

Section 230 (c) of the Communications Decency Act (1996)

- Protection for “Good Samaritan” blocking and screening of offensive material
 - Platforms are not **publishers or speakers**, thus not liable for user content (c)(1)
 - **Users not liable** for reposting other user’s content
 - Platforms also protected if they **remove objectionable content** (c)(2)

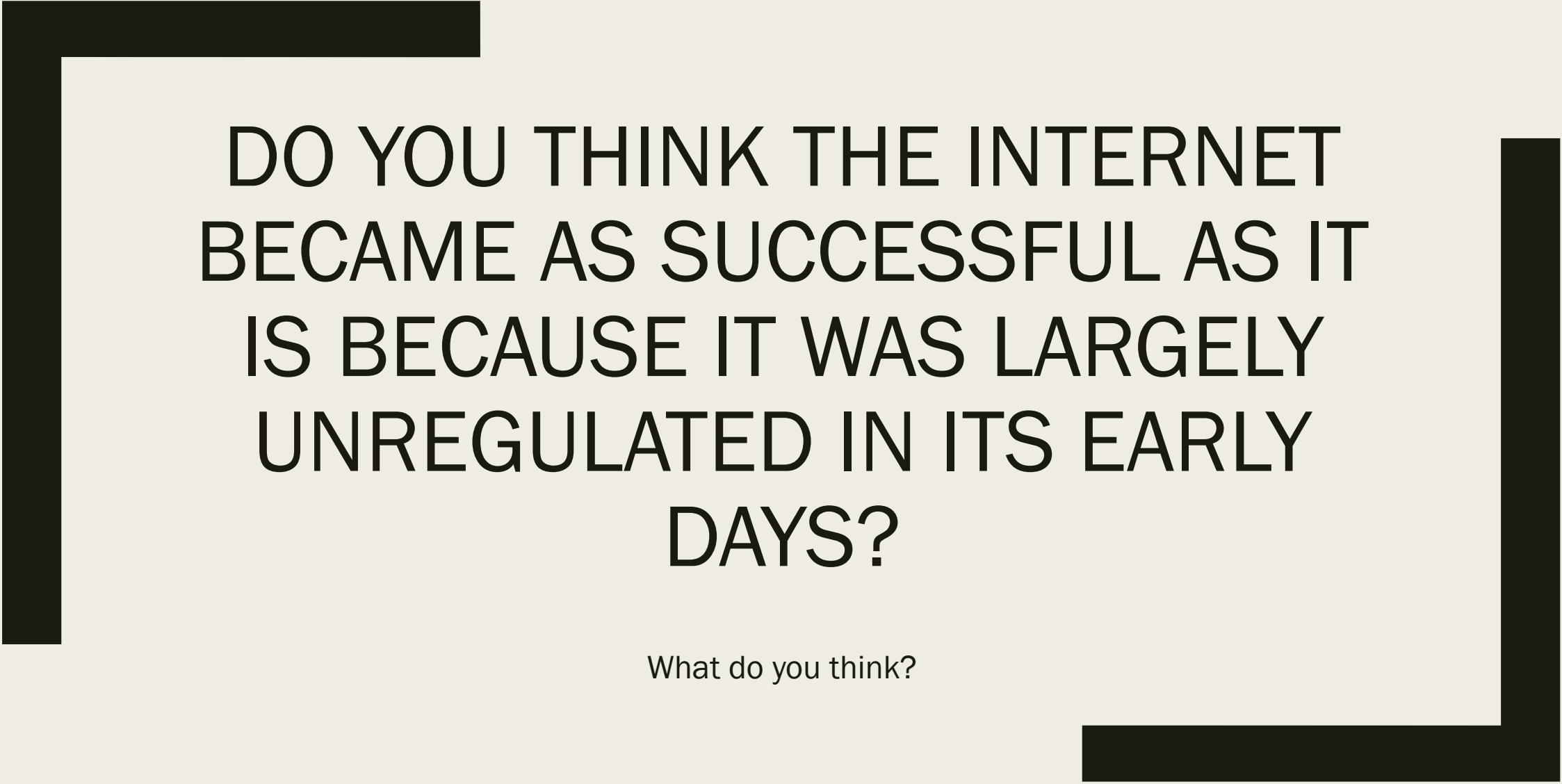
The origin of current Section 230 doctrine

Zeran v. America (1997)



"Section 230 was enacted, in part, to maintain the robust nature of Internet communication and, accordingly, to keep government interference in the medium to a minimum.

Congress recognized the Internet and interactive computer services as offering "a forum for a true diversity of political discourse, unique opportunities for cultural development, and myriad avenues for intellectual activity. It also found that the Internet and interactive computer services "have flourished, to the benefit of all Americans, with a minimum of government regulation."



DO YOU THINK THE INTERNET
BECAME AS SUCCESSFUL AS IT
IS BECAUSE IT WAS LARGELY
UNREGULATED IN ITS EARLY
DAYS?

What do you think?

Limits to platform's immunity

- despite the influence of *Zeran*, there have been decisions that have avoided granting immunity to interactive computer services. These have stemmed largely from two theories:
 - that the sites have contributed to illegal content
 - that the sites were engaged in activities other than publishing and speaking.

US rules for content moderation: Key takeaways

- Basically **no punishment for tellings lies or engaging in hate speech**
- Almost Immunity for **online platforms and users** when disseminating third party content
- the role of content moderation is left to the **market**. Remember that since this is federal law, states cannot enact any legislation that goes against it.

As a result:

- “Open microphone” business models (FB, Youtube, X), were enabled by and rely on the Section 230’s shelter. If and as the protection changes, so very likely will the business models.
- These same business models fostered hate speech (online bullying), illegal content (child pornography, revenge porn) and disinformation (fake news)
- The market – users and advertisers (and other hidden forces?) – drives platform’s moderation choice.

The U.S. system leads the world's approach to content moderation

- U.S. free-speech culture exported worldwide: European law increasingly “Americanized” on hate-speech limits.
- Most major platforms are US-based: the role of anchoring regulation

ARE THESE RULES STILL
FIT FOR TODAY'S DIGITAL
WORLD?



The world has changed a lot in 20 years

- **Traditional news sources have lost credibility;** many people get information from social media, influencers, friends, and family.
- We face an **information overload**, and **falsehoods often appear credible**.
- Social networks are used strategically to spread disinformation and even to disrupt elections.
- We now understand **behavioral biases** — confirmation bias, dopamine rewards; anger-inducing content, echo-chambers, etc. drive engagement.
- **Algorithms**, not humans, moderate most content. They are programmed to maximize engagement
- **Growing criticism** against 1st Amendment and Section 230:
 - They have given tech companies too much power, stifling competition and innovation.
 - They have allowed harmful content to thrive on social media.

The world is worried

Global Risks Report 2024

Top 10 risks



"Please estimate the likely impact (severity) of the following risks over a 2-year and 10-year period."

2 years



10 years



Risk categories | Economic | Environmental | Geopolitical | Societal | Technological

Source: World Economic Forum Global Risks Perception Survey 2023-2024.

The platform's response: Meta's shift



2018

*"We're **proactively enforcing our policies to remove more harmful content**, preventing borderline content from spreading, giving people more control of their experience, and **creating independent oversight and transparency into our systems**"*

The platform's response: Meta's shift



2025


"The recent elections feel like a cultural tipping point towards once again prioritizing speech

Starting in the US, we are ending our third-party fact-checking program and moving to a Community Notes model, similar to X.

We will get rid of restrictions on topics like immigration and gender (hate speech) because they are part of mainstream discourse

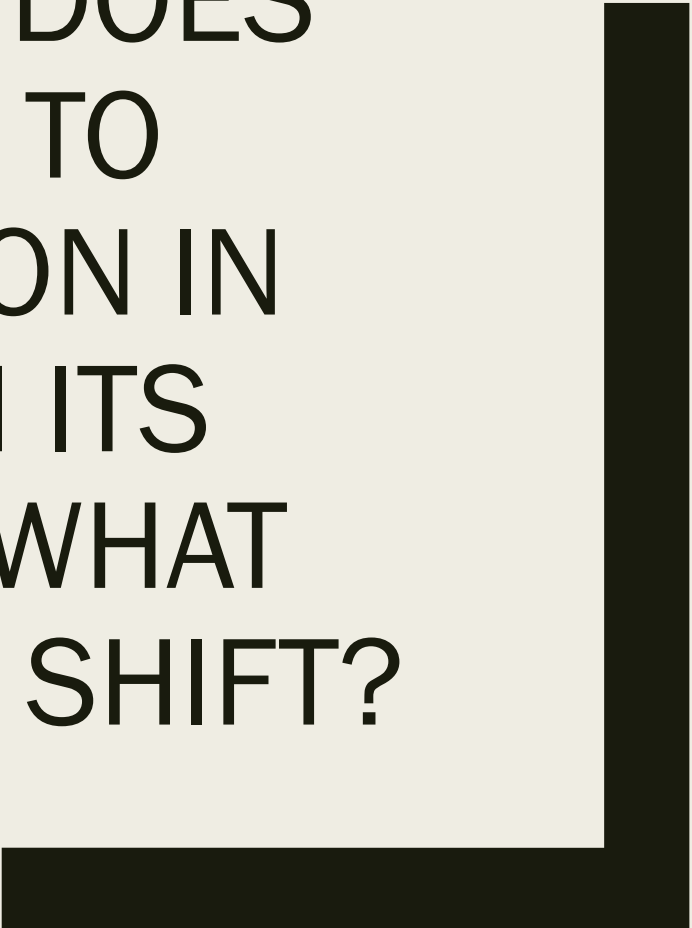
We will eliminate our systems to detect minor violations of content moderation policies.

We are going to work with president trump to push back on governments around the world going after American companies".



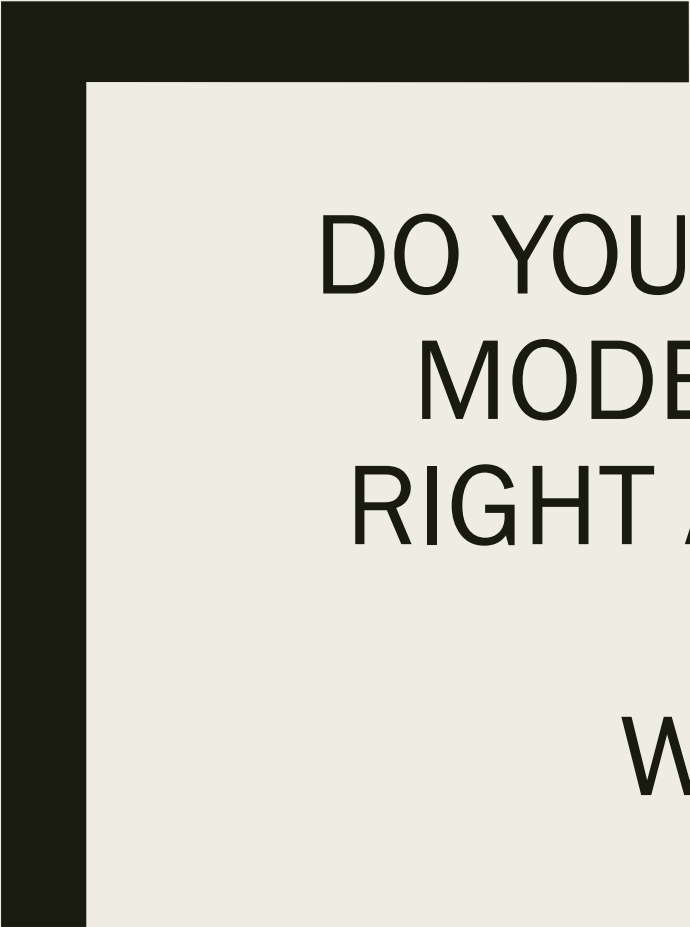
IN A NUTSHELL, HOW DOES
META'S APPROACH TO
CONTENT MODERATION IN
2025 DIFFER FROM ITS
APPROACH IN 2018, WHAT
FACTORS EXPLAIN THIS SHIFT?

What do you think?



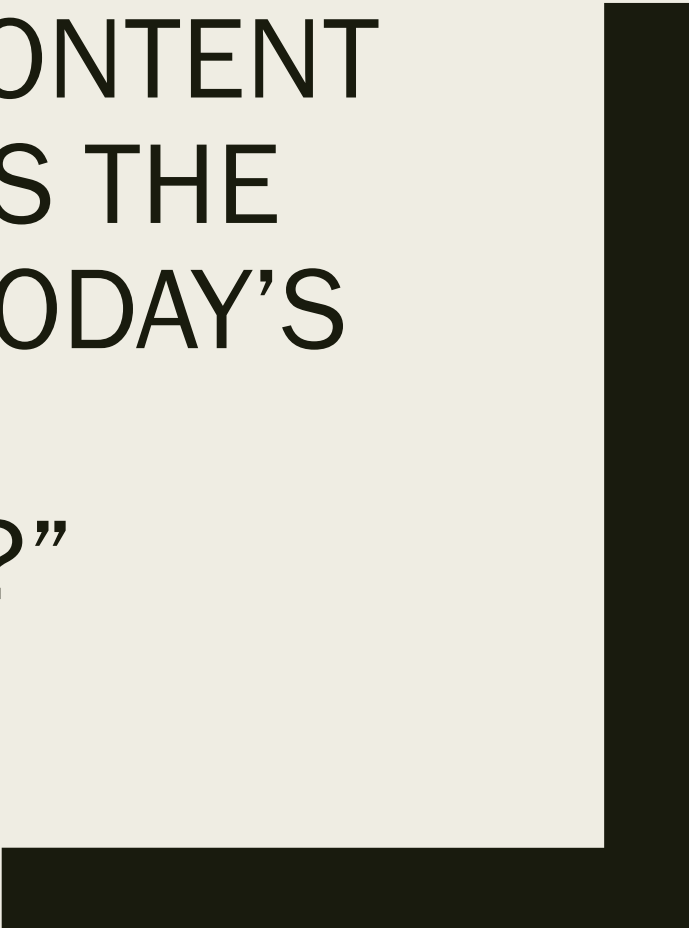
Possible changes to the current system coming soon?

- Congress: Divided, slightly Republican → unlikely to reform Section 230
- Supreme Court: Avoiding direct rulings on Section 230 and the First Amendment
 - Section 230 – 2023: Cases on whether recommendation algorithms amplifying terrorism-related content lose immunity
 - Ruling: Families' claims rejected under the Antiterrorism Act → no comment on Section 230
 - First Amendment – 2024: Cases on whether states can restrict platforms' content-moderation practices
 - Ruling: State laws were too broad → could over-restrict speech
 - Platforms have First Amendment rights—like newspapers or bookstores—to choose and organize content (*but... weren't platforms specifically not publishers or speakers under section 230?*)



DO YOU THINK THE US CONTENT
MODERATION MODEL IS THE
RIGHT APPROACH FOR TODAY'S
DIGITAL WORLD?
WHY OR WHY NOT?"

What do you think?



THANK YOU FOR YOUR ATTENTION!

Questions?

[Bea.gquiroga@gmail.com.](mailto:Bea.gquiroga@gmail.com)

<https://www.linkedin.com/in/beatriz-garcia-quiroga/>