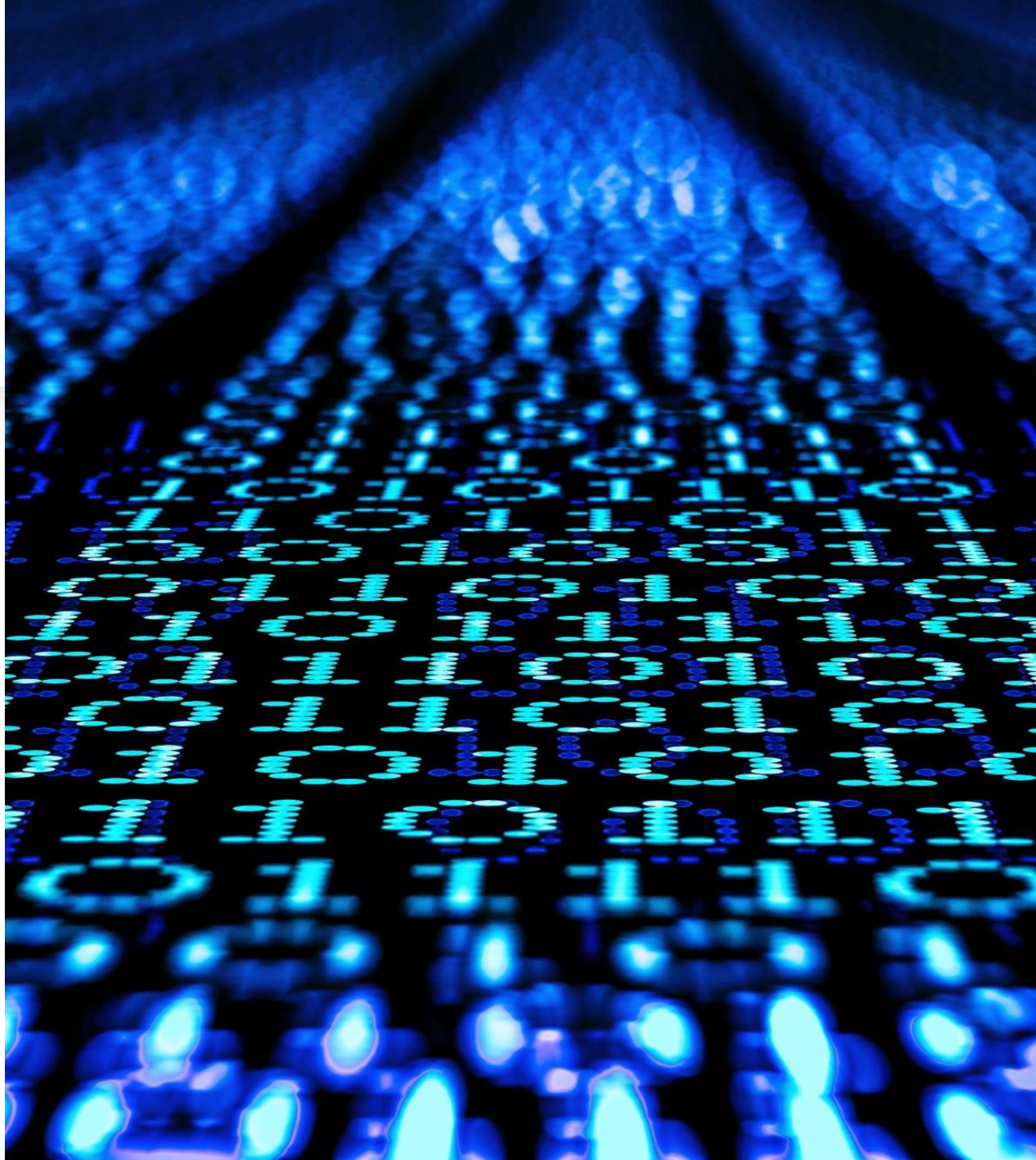
“Hate speech, disinformazione e responsabilità”

Dott.ssa Katia De Blasio
Università Roma Tre
28 marzo 2025



LA NASCITA DI INTERNET

- Negli anni ‘90 del secolo scorso, sia negli USA, sia in Europa, ove si voleva favorire lo sviluppo del mercato interno, si tendeva a considerare la regolamentazione come un ostacolo all’innovazione e allo sviluppo economico tramite *Internet*;
- Negli USA tale approccio veniva consolidato attraverso il *Communication Decency Act*→ si introduceva un sistema di esenzione della responsabilità per il quale gli intermediari, ossia le piattaforme, venivano considerati estranei rispetto alla condotte illecite di terzi nell’utilizzo della piattaforma stessa (*safe harbour* e *good samaritan clause*); Centralità del *First Amendment* della Costituzione americana nel riconoscimento di queste immunità.

COMMUNICATION DECENCY ACT (1996), § 230 PROTECTION FOR PRIVATE BLOCKING AND SCREENING OF OFFENSIVE MATERIAL

[…] (a) Findings

The Congress finds the following:

(1) The rapidly developing array of Internet and other interactive computer services available to individual Americans represent an extraordinary advance in the availability of educational and informational resources to our citizens […]

(4) The Internet and other interactive computer services have flourished, to the benefit of all Americans, with a ***minimum of government regulation.***

(c) Protection for "Good Samaritan" blocking and screening of offensive material

- **(1)** ***Treatment of publisher or speaker***

No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider.

- **(2)** ***Civil liability***

No provider or user of an interactive computer service shall be held liable on account of—

(A) any action voluntarily taken in good faith to restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected; or

(B) any action taken to enable or make available to information content providers or others the technical means to restrict access to material described in paragraph (1)

Eccezioni se l’azione è fondata su:
-diritto penale;
-proprietà intellettuale;
-leggi statali che siano coerenti coi principi della §230;
-Electronic Communications Privacy Act (1986);
-Fight Online Sex Trafficking Act of 2017 (FOSTA).

TEST APPLICATO DALLE CORTI STATUNITENSI:

1. Il convenuto è il fornitore di un servizio internet interattivo? Is the defendant a provider or user of an interactive computer service?

2. L'attore invoca la responsabilità del convenuto quale editore? Does the plaintiff seek to hold the defendant liable as a publisher or speaker?

3. L'azione dell'attore trova origine nelle informazioni fornite da un terzo? Does the plaintiff's claim arise from information provided by another information content provider?

→ Se la risposta è ***sì*** per tutte e tre i quesiti, allora la responsabilità del *service provider* è esclusa.

Le corti statunitensi hanno interpretato in maniera estensiva la §230 CDA, riconoscendo, in sostanza, l'immunità da responsabilità civile ad ogni sito *web* o servizio *online* ospitante contenuti generati dagli utenti (in *Zeran v. America Online, Inc.* 129 F.3d 327 (1997), si sostiene che la §230 esclude l'accoglimento delle "lawsuits seeking to hold a service provider liable for its exercise of a publisher's traditional editorial functions—such as deciding whether to publish, withdraw, postpone or alter content".

la § 230 CDA è stata definita come «***twenty-six words that created the internet***»

(J. Kosseff, *The Twenty-Six Words That Created the Internet*, Cornell University Press, 2019)

Social media e moderazione dei contenuti

- Per i *social media,* la §230 è stata interpretata nel senso di escludere la responsabilità delle piattaforme nelle cause basate sulla loro decisione di rendere visibili o rimuovere contenuti generati dagli utenti.
- Ci si può chiedere se l'utilizzo di algoritmi per filtrare ed organizzare i contenuti possa ricadere al di sotto della §230 CDA o meno. In alcuni casi giudiziari, gli attori hanno provato a sostenere che in questi casi i *service provider* fossero coinvolti nella creazione di contenuto di terze parti o che comunque questa attività non fosse di tipo editoriale (non un *publisher,* ma un *distributor*), ma le corti sino ad ora hanno rigettato queste teorie (*Force v. Facebook*, Inc., 934 F.3d 53, 66–69 (2d Cir. 2019), ove gli attori erano stati vittime di un attacco terroristico facilitato dal coordinamento dei terroristi su Facebook. In particolare la Corte ha ritenuto che il come e il dove fornire il contenuto sia una scelta editoriale. Inoltre per il primo argomento, la Corte ha ritenuto che Facebook non partecipasse alla creazione di informazioni per il semplice fatto di mostrarle a determinati utenti senza alterarne il contenuto.
- L'algoritmo, in *Force*, è considerato un "*neutral tool*".

First Amendment della Costituzione Americana

→ Protegge il ***freedom of speech*** in varie declinazioni. Negli ambienti *online:*

-Protegge la libertà di espressione degli utenti di siti *web* e piattaforme;

-Protegge la libertà del *web designer* nel *design* del sito *web*;

-Protegge la libertà di decidere sulla pubblicazione/censura di contenuti di terze parti.

- Uno dei principi fondanti il costituzionalismo americano è quello per il quale la Costituzione limiti il diritto pubblico, ma non quello privato; dunque il diritto costituzionale interviene se vi è una «*state* action» (cfr. caso *Civil rights cases*, 109 U.S. 3 (1883). Pertanto, il primo emendamento trova applicazione in caso di censura da parte dell'attore pubblico (si veda il caso di Trump che bloccava i dissidenti sui suoi profili social, *Knight First Amendment Ins. At Columbia Univ. v. Trump* -928 F. 3d 226 (2° Cir. 2019), ma non inficia la libertà delle piattaforme *online* nel prendere decisioni autonome in tema di moderazione dei contenuti condivisi e la libera espressione degli utenti. La questione è poi finita davanti alla Corte Suprema dopo la vittoria di Biden alle elezioni e la sentenza della corte d'appello è stata annullata in quanto intentata nei confronti di Trump quale presidente. Il Justice Thomas ha depositato una *concurring opinion* ove sostiene che le piattaforme *online* possono essere esentate da responsabilità estendendo la disciplina sui *common carriers* (vettori comuni) o quella concernente le *public accomodations.*

- I *social media* in sé non esercitano funzioni pubbliche né possono essere considerati *State actors*, si veda il caso *Prager University v. Google LLC*, No. 18-15712, 2020). Nel caso di specie la *Prager University* aveva citato in giudizio YouTube asserendo che la piattaforma, in violazione del *First amendment*, avesse applicato il *tag* di «restricted mode» ad alcuni video pubblicati dall'istituzione ed avesse ingiustamente tolto la monetizzazione ai video. La *District court* e la *Court of Appeals* del nono circuito hanno rigettato la domanda sostenendo che il primo emendamento non trovasse applicazione poiché YouTube è un attore privato, e «The Free Speech Clause of the First Amendment prohibits the government—not a private party—from abridging speech».

IMPORTAZIONE DEL MODELLO STATUNITENSE IN EUROPA

- Gli artt. 12,13 e 14 della Direttiva sul commercio elettronico (Direttiva 2000/31/CE) distinguevano tra:

-*Semplice trasporto* (consistente nel trasmettere, su una rete di comunicazione, informazioni fornite da un destinatario del servizio, o nel fornire un accesso alla rete di comunicazione);

-*Caching* (memorizzazione automatica, intermedia e temporanea di informazioni effettuata al solo scopo di rendere più efficace il successivo inoltro ad altri destinatari);

-*Hosting* (memorizzazione *non temporanea* di informazioni fornite da un destinatario del servizio)

In entrambi i casi, il *service provider* non era responsabile per l'illiceità delle suddette informazioni purché sussistessero determinate condizioni.

In particolare egli doveva provvedere a rimuovere suddette informazioni se un organo giurisdizionale o un'autorità amministrativa ne avesse disposto la rimozione o la disabilitazione dell'accesso (*caching*) o appena fosse venuto a conoscenza dell'illiceità delle informazioni (*hosting*).

Art. 15, co. 1 Assenza dell'obbligo generale di sorveglianza

«Nella prestazione dei servizi di cui agli articoli 12, 13 e 14, gli Stati membri ***non impongono ai prestatori un obbligo generale di sorveglianza*** sulle informazioni che trasmettono o memorizzano né un obbligo generale di ricercare attivamente fatti o circostanze che indichino la presenza di attività illecite».

CGUE SABAM c. Netlog, causa C-360/10. La SABAM è la società belga che gestisce i diritti d'autore di musicisti e ha il compito di autorizzare l'utilizzo da parte di terzi delle loro opere. Netlog permetteva ai propri utenti di inserire della musica nella loro pagina personale. SABAM chiedeva dunque a Netlog di impegnarsi a cessare *immediatamente e per il futuro* la messa a disposizione del pubblico non autorizzata, ma Netlog sosteneva che ciò avrebbe imposto sulla piattaforma un obbligo generale di sorveglianza *ex* art 15 della Dir. Nel bilanciare la tutela del diritto d'autore e la libertà d'impresa dell'*hosting provider,* la CGUE ha ritenuto che l'implementazione di un tale sistema di filtraggio implicasse una sorveglianza sulla totalità o sulla maggior parte delle informazioni memorizzate presso il prestatore di servizi di hosting coinvolto e fosse troppo onerosa per Netlog (oltre a violare la *privacy* degli utenti).

Decisioni nella stessa direzione: *Glawischnig-Piesczek v. Facebook Ireland* Ltd, causa C-18/18, *Scarlet v. SABAM*, C-70/10

- Questo regime poteva essere condivisibile qualche decennio fa, ma negli sviluppi più recenti emergono nuovi modelli di *business* in cui l'organizzazione dei contenuti e la moderazione avvengono con l'ausilio di strumenti di intelligenza artificiale. La moderazione dei contenuti è essenziale per gli scopi di profitto delle aziende in quanto volta a costruire uno spazio digitale nel quale offrire annunci pubblicitari, dato che la pubblicità *online* è la fonte di profitto principale per le piattaforme digitali → si mette in discussione il regime di irresponsabilità delle piattaforme;
- Contributo delle piattaforme *online* nel genocidio in Myanmar, scandalo Cambridge analytica e influenza di potenze straniere sulle elezioni in Europa e Stati Uniti, ecc…;
- Gli intermediari gestiscono spazi che sono a cavallo tra il pubblico e il privato, tanto che spesso gli è stata demandata la corresponsione di informazioni a pubblici poteri (non a caso vengono definiti *gatekeepers*);

CAPO II DIGITAL SERVICES ACT (Regolamento (UE) 2022/2065, «DSA»)

- Dagli **artt. 4 all'8** si mantengono le regole generali in tema di irresponsabilità dei prestatori di servizi per le condotte degli utenti;
- In particolare, dagli artt. 4 al 6 si mantengono le esenzioni di responsabilità in caso di semplice trasporto, *caching* e *hosting* di informazioni prodotte da terzi;
- All'art. 7 si prevede l'esenzione dalla responsabilità nel caso in cui i prestatori di servizi svolgano «***in buona fede*** e in modo diligente, **indagini volontarie** di propria iniziativa» o decidano «di adottare altre misure volte a **individuare, identificare e rimuovere contenuti illegali** o a disabilitare l'accesso agli stessi [...]»
- All'art. 8 si ribadiscono le prescrizioni del vecchio art. 15 della Direttiva sul commercio elettronico, per cui «ai prestatori di servizi intermediari **non è imposto alcun obbligo generale di sorveglianza** sulle informazioni che tali prestatori trasmettono o memorizzano, né di accertare attivamente fatti o circostanze che indichino la presenza di attività illegali».

MECCANISMI DI TRASPARENZA E DI SEGNALAZIONE

- Art. 14 = **Termini e condizioni** «I prestatori di servizi intermediari includono nelle loro condizioni generali informazioni sulle restrizioni che impongono in relazione all'uso dei loro servizi per quanto riguarda le informazioni fornite dai destinatari del servizio. Tali informazioni riguardano tra l'altro **le politiche, le procedure, le misure e gli strumenti utilizzati ai fini della moderazione dei contenuti**, compresi il processo decisionale algoritmico e la verifica umana, nonché le regole procedurali del loro sistema interno di gestione dei reclami»;

- Art. 15 = **Obblighi in materia di relazioni di trasparenza per i prestatori di servizi intermediari=** I prestatori di servizi intermediari mettono a disposizione del pubblico, in un formato leggibile meccanicamente e in modo facilmente accessibile, almeno una volta all'anno, relazioni chiare e facilmente comprensibili sulle **attività di moderazione dei contenuti** svolte durante il periodo di riferimento.

- Art 16 = **Meccanismo di segnalazione e azione= «**I prestatori di servizi di memorizzazione di informazioni predispongono meccanismi per consentire a qualsiasi persona o ente di notificare loro la presenza nel loro servizio di informazioni specifiche che tale persona o ente ritiene costituiscano contenuti illegali».

- Art. 17 = **Motivazione** «I prestatori di servizi di memorizzazione di informazioni forniscono a tutti i destinatari del servizio interessati una motivazione chiara e specifica per le seguenti restrizioni imposte a motivo del fatto che le informazioni fornite dal destinatario del servizio costituiscono contenuti illegali o sono incompatibili con le proprie condizioni generali: a) eventuali restrizioni alla visibilità di informazioni specifiche fornite dal destinatario del servizio, comprese la rimozione di contenuti, la disabilitazione dell'accesso ai contenuti o la retrocessione dei contenuti […]»

OBBLIGHI SUPPLEMENTARI A CARICO DEI FORNITORI DI PIATTAFORME ONLINE DI DIMENSIONI MOLTO GRANDI E DI MOTORI DI RICERCA ONLINE DI DIMENSIONI MOLTO GRANDI PER LA GESTIONE DEI RISCHI SISTEMICI

- Art. 34= Le piattaforme di determinate dimensioni debbono effettuare, annualmente, delle valutazioni **«dei rischi sistemici** derivanti dalla progettazione, compresi sistemi algoritmici, dal funzionamento e dall'utilizzo dei loro servizi nell'Unione», quali, ad esempio «**la diffusione di contenuti illegali tramite i loro servizi**», «eventuali effetti negativi, attuali o prevedibili, per l'esercizio dei diritti fondamentali», «eventuali **effetti negativi, attuali o prevedibili, sul dibattito civico e sui processi elettorali, nonché sulla sicurezza pubblica**», «qualsiasi effetto negativo, attuale o prevedibile, in relazione alla violenza di genere, alla protezione della salute pubblica e dei minori e alle gravi conseguenze negative per il benessere fisico e mentale della persona».
- Nello svolgimento della valutazione dei rischi *ex* art. 34, le piattaforme devono tenere in considerazione diversi fattori, tra cui anche la progettazione del loro **sistema di raccomandazione** e i loro **sistemi di moderazione dei contenuti**

- Art. 35 fornisce degli esempi di misure che le grandi piattaforme possono adottare per attenuare i rischi (ad es. l’adeguamento della progettazione, delle caratteristiche o del funzionamento dei loro servizi, anche delle loro interfacce online o l'avvio o l'adeguamento della cooperazione con altri fornitori di piattaforme online o di motori di ricerca online attraverso i codici di condotta, ecc…);
- Art 36= In caso di crisi (circostanze eccezionali comportano una grave minaccia per la sicurezza pubblica o la salute pubblica nell'Unione), la Commissione, su raccomandazione del comitato, può adottare una **decisione** che impone a una o più piattaforme online di: valutare il funzionamento e l'uso dei loro servizi contribuiscano, o possano contribuire, in maniera significativa a una minaccia grave; adottare una delle misure di cui all’art. 35; redigere una relazione concernente l’impatto della piattaforma sulla minaccia grave e le conseguenti misure adottate;
- Art. 37= Le grandi piattaforme devono sottoporsi, almeno una volta all’anno e a loro spese, a revisioni indipendenti volte a verificare il rispetto degli obblighi del capo II ed eventualmente agli impegni assunti a norma dei codici di condotta volontari;
- Art. 40 = «potere ispettivo» volto a vigilare sul rispetto del regolamento riconosciuto alla Commissione, al coordinatore dei servizi digitali del luogo di stabilimento e ad un gruppo di «ricercatori abilitati»

CODICE DI BUONE PRATICHE SULLA DISINFORMAZIONE

- Istituito nel 2018 e rafforzato nel 2022. In data 13 febbraio 2025 la Commissione e il comitato europeo per i servizi digitali hanno approvato l'integrazione del codice di buone pratiche sulla disinformazione 2022 come codice di condotta sulla disinformazione **nel quadro della legge sui servizi digitali**. In quanto tale, il codice diventerà un parametro di riferimento pertinente per determinare la conformità alla legge sui servizi digitali per quanto riguarda i rischi di disinformazione per i fornitori di piattaforme online di dimensioni molto grandi e di piattaforme online di dimensioni molto grandi che aderiscono ai suoi impegni e li rispettano;
- L’adesione al codice di condotta può costituire una misura ragionevole di attenuazione dei rischi per le piattaforme di grandi dimensioni (cfr. art 45);
- Contiene 44 impegni e 128 misure specifiche nei seguenti settori: demonetizzazione la disinformazione; trasparenza della pubblicità politica; rafforzare la comunità di fact-checking;

SANZIONI

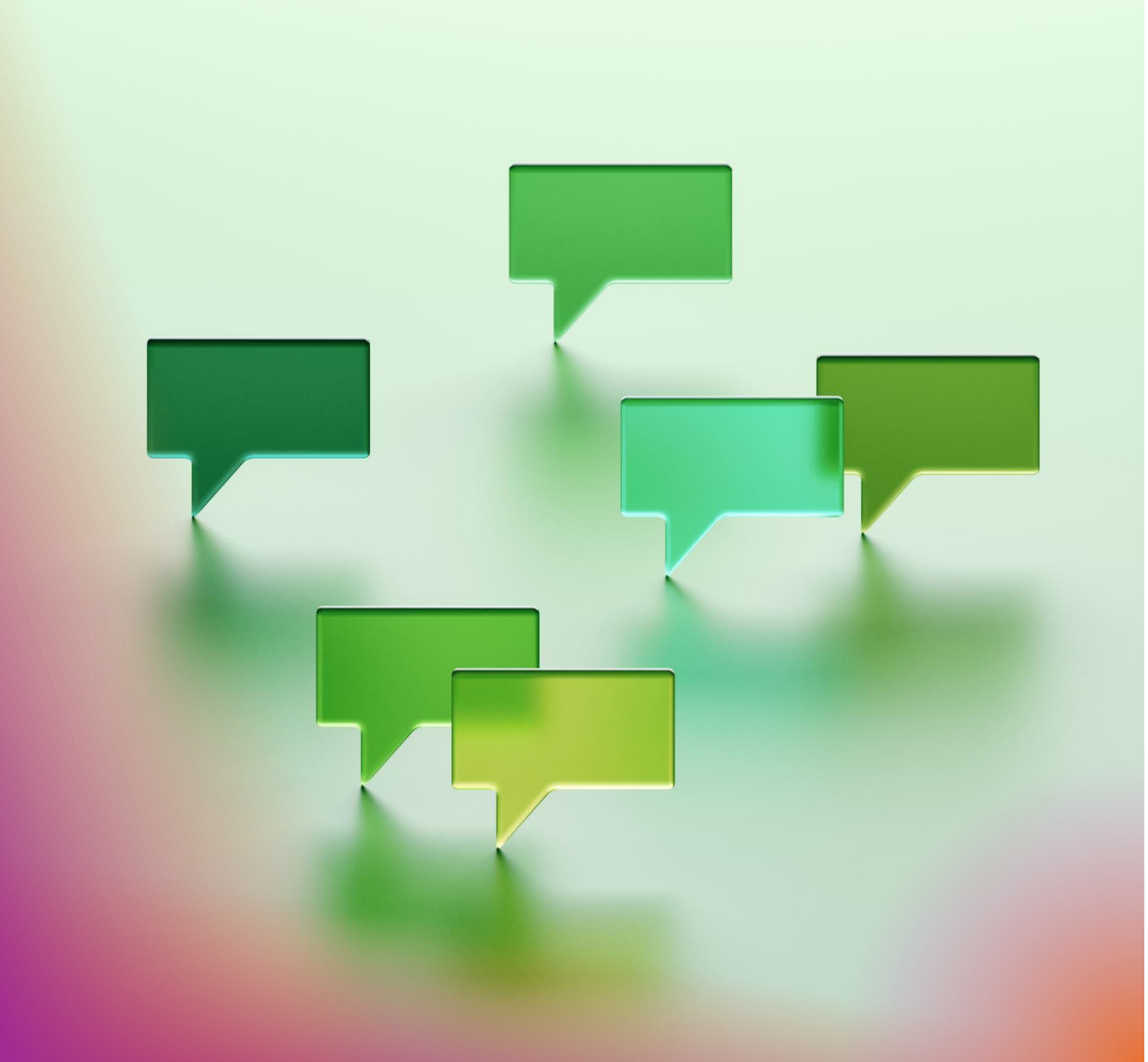
Gli stati membri stabiliscono le norme relative alle sanzioni applicabili in caso di violazione del DSA da parte dei fornitori di servizi intermediari che rientrano nella loro competenza;

L’importo massimo delle sanzioni pecuniarie che possono essere irrogate in caso di inosservanza di un obbligo stabilito dal DSA è pari al 6% del fatturato annuo mondiale del fornitore di servizi (art. 52)

Comparazione con gli Stati Uniti

- Negli Stati Uniti la disciplina del §230 è rimasta immutata nonostante nel corso degli anni siano stati introdotti diversi *bill* con l'intento di modificarla;
- TUTTAVIA, un' interessante giurisprudenza recente ha cominciato a riconoscere la responsabilità del produttore in capo al *service provider*

A.M. v. Omegle.com LLC, 3:21-cv-01674-MO, Omegle era un *social media* che metteva in contatto sconosciuti. Nel caso di specie una bambina di 11 anni era stata messa in contatto con un individuo che aveva poi commesso degli abusi sessuali nei suoi confronti. Il giudice ha ritenuto che la piattaforma Omegle avesse dovuto essere concepita in modo differente. Il fatto che non richiedesse la verifica dell'età dell'utente, la pubblicità che invogliava gli utenti a «parlare con gli sconosciuti!» rivolta a potenziali utenti minorenni e l'assenza di meccanismi volti a verificare l'identità e l'età degli utenti, e a separare i minorenni dai maggiorenni, rendevano la piattaforma difettosa.



Poche settimane fa, Meta ha annunciato (al momento solamente per gli USA) l'introduzione di nuove modalità di *content moderation* per Instagram e Facebook al fine di sostituire i *fact checkers* con un meccanismo di «*community notes*» analogo a quello già implementato nel *social network* X, sostenendo che tale novità consentirà di meglio garantire la libertà di espressione, mentre il sistema precedente portava a delle ingiuste «censure» (https://about.fb.com/news/2025/01/meta-more-speech-fewer-mistakes/)

Problematiche nelle recenti elezioni in Romania

Lo scorso 17 dicembre la Commissione europea ha aperto un'indagine nei confronti di Tik Tok (https://ec.europa.eu/commission/presscorner/detail/en/ip_24_6487) concernente presunte violazioni degli articoli 34(1), 34(2) e 35(1) del DSA.
L'indagine sarà incentrata sulla gestione del rischio nel dibattito politico per le elezioni, con riferimento a:

- Termini e condizioni per la fruizione di pubblicità politica sulla piattaforma;
- Sistema di raccomandazione dei contenuti di TikTok.

Grazie dell’attenzione!

katia.deblasio@uniroma3.it