

Disinformation and Hate Speech Online: From self regulation to Co/Hard Regulation within the European Digital Single Market?

Prof. Oreste Pollicino

Bocconi University



@orestepollicino

Constitutional Transatlantic Conflicts

Liberty

vs

Dignity

The European constitutional matrix

- Right to be informed (if not in a veritable way, in a verifiable way)
- hate speech has a constitutional relevance *per se*

Disinformation

- **(US) Free market place of ideas**

VS

- **(digital) distrust versus Pareto efficiency in the information market**

Hate speech

European Court of Human Rights, 6-7-2006

Brandenburg v. Ohio, 395 U.S. 444 (1969)

the imminent lawless action test

ECHR (2006) “[T]olerance and respect for the equal dignity of all human beings constitute the foundations of a democratic, pluralistic society. That being so, as a matter of principle it may be considered necessary in **certain democratic societies to sanction or even prevent all** forms of expression which spread, incite, promote or justify hatred based on intolerance ..., provided that any ‘formalities’, ‘conditions’, ‘restrictions’ or ‘penalties’ imposed are **proportionate** to the legitimate aim pursued”

Transatlantic Clash

- Digital Platforms based in USA, allegedly regulated by US Constitutional values

But

- Providing services in the European Digital Single Market

Sounding Board Unanimous Final Opinion on the Code of practice on Disinformation

26-4-2018

“The so called -Code of practice - as presented by the working group contains no common approach, no clear and meaningful commitments, no measurable objectives or KPIs, hence no possibility to monitor process, and no compliance or enforcement tool: it is by no means self-regulation, and therefore the Platforms, despite their efforts, have not delivered a Code of Practice”

Facebook plans launch of its own “Supreme Court” for handling takedown appeals

Privatization of digital justice ?

- If someone disagrees with a decision we've made, they can appeal to us first, and soon they will be able to further appeal this **to the independent board...** As an independent organization, we hope it gives people confidence that their views will be heard and that Facebook doesn't have the ultimate power over their expression.“

Mark Zuckerberg

Status quo

- **Failure and (consequent) European scepticism for pure self-regulation**

The Judicial path

- Judicial activism reacting to obsolence of the legal regime and legislative inertia
- A jeopardised fragmented approach
- Active/passive
- What about principle of legal certainty ?

Active and passive providers?

Italian Supreme Court 7708/2019

- Going beyond the decision of the Milan Court of Appela in *RTI v. Yahoo!* (2015) according to which all hosting providers benefit from the exemption of liability if they are not directly aware of the uploaded content of the user.
- The Supreme Court clarified the relevance of the distinction between active and passive hosting providers
- Active hosting when providers filter, select, index organize, catalogue, aggregate, rate, use, edit, extract, promote online content

24-9-2019

Google. inc and the European boundaries of the Right to Be Forgotten

- Private Arbiter between the clash of fundamental rights
- Delegating public functions to private operators
- Geographic limitations

In the meantime...

Glawischnig-Piesczek - C-18/18

The Directive on electronic commerce **does not preclude a court of a Member State from ordering a host provider:**

- a worldwide removal

Glawischnig-Piesczek - C-18/18

- The Directive **does not preclude a court of a Member State from ordering a host provider to** remove information which it stores, the content of which is **equivalent** to the content of information which was previously declared to be unlawful, or to block access to that information (...) **provided that the differences in the wording of that equivalent content, compared with the wording characterizing the information which was previously declared to be illegal, are not such as to require the host provider to carry out an independent assessment of that content (thus, the host provider may have recourse to automated search tools and technologies)**

**Thomas Hughes , executive director of the
charity Article 19, ‘**

“This judgment has major implications for online freedom of expression around the world”

- **Compelling social media platforms like Facebook to automatically remove posts regardless of their context will infringe our right to free speech and restrict the information we see online.**
- **The ruling also means that a court in one EU member state will be able to order the removal of social media posts in other countries, even if they are not considered unlawful there. This would set a dangerous precedent where the courts of one country can control what internet users in another country can see”**

The EU strategy for disinformation and hate speech online

Self-Regulation

Co-regulation

Hard Regulation

A Self-Regulatory Attempt?

- Why a self-regulatory attempt for addressing hate speech and disinformation online?
 - Mitigating the risk of collateral censorship
 - Flexibility in reviewing and monitoring improvements
 - Complexitiy in defining legally disinformation and hate speech
 - But...some drawbacks, especially legal certainty due to platforms' discretion in establishing their standards of prote

The disinformation framework

Communication - Tackling online disinformation: a European Approach

High-Level Expert Group on Disinformation

Code of practice on disinformation

AGCOM: a table on online disinformation

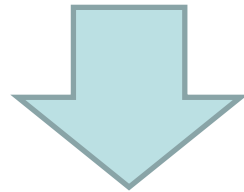
- The Italian Regulatory Authority of Communication (AGCOM) has set up a technical group on online disinformation established by Resolution No. 423/17/CONS.
- The aim is to create an institutional forum in order to encourage self-regulatory solutions and the exchange of good practices for identifying and contrasting online disinformation.
- At the meeting are present Google, Facebook, Wikipedia and representatives of the major press and broadcasting groups and their respective trade associations.

Code of practice on hate speech

- Code of conduct on countering illegal hate speech online adopted on 31 May 2016
- The Code was developed following the October 2015 EU Colloquium on Fundamental Rights on “Tolerance and respect: preventing and combating anti-semitic and anti-muslim hatred in Europe”, and the December 2015 EU Internet Forum.
- Signatories: Facebook, Google, Microsoft, Twitter, Instagram, Google+, Snapchat, Dailymotion and jeuxvideo.com
- However, in recent months, some European countries (such as Germany, the UK, France and Ireland) stated their intentions to introduce regulation concerning hate speech

The EU strategy for disinformation and hate speech online

Failure of self-regulation and co-regulation



Hard Regulation

From Napster to tech giants...



...from the e-Commerce Directive via the Digital Single Market Strategy to the Digital Service Act?



A new approach to platform regulation?

Communication on Online Platforms and the Digital Single Market Opportunities and Challenges for Europe - COM 2016 (288)

- **a level playing field for comparable digital services;**
- **responsible behaviour of online platforms to protect core values;**
- transparency and fairness for maintaining user trust and safeguarding innovation;
- open and non-discriminatory markets in a data-driven economy.

A new approach to platform regulation?

Commission Recommendation of 1.3.2018 on measures to effectively tackle illegal content online - C(2018) 1177

- Submitting and processing notices
- Informing content providers and counter-notices
- Out-of-court dispute settlement
- Transparency
- Proactive measures

Reviewing the e-Commerce Directive?

The system of the e-Commerce Directive has not been repealed

BUT

Directive 2019/790 - Copyright Directive

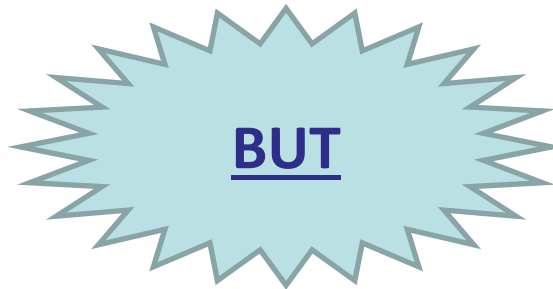
Regulation on Terrorism Content (Proposal)

Directive 2018/108 –AVMS Directive

Copyright Directive

Act of communication to the public – Art. 17

When an online content-sharing service provider performs an act of communication to the public or an act of making available to the public, the limitation of liability established in Article 14(1) of Directive 2000/31/EC shall not apply to the situations covered by this Article.



Article 14(1) of Directive 2000/31/EC is still applicable to those service providers for purposes falling outside the scope of this Directive

Copyright Directive

A new exception of liability? – Art. 17

If no authorisation is granted, online content-sharing service providers shall be liable for unauthorised acts of communication to the public, including making available to the public, of copyright-protected works and other subject matter, **unless the service providers demonstrate that they have:**

(a) made best efforts to obtain an authorisation, and

(b) made, in accordance with high industry standards of professional diligence, best efforts to ensure the unavailability of specific works and other subject matter for which the rightholders have provided the service providers with the relevant and necessary information; and in any event

(c) acted expeditiously, upon receiving a sufficiently substantiated notice from the rightholders, to disable access to, or to remove from their websites, the notified works or other subject matter, and made best efforts to prevent their future uploads in accordance with point (b).

New responsibilities also from data protection law?

Rise of Digital Privacy

(Digital Rights Ireland, Google Spain, Schrems)

It has **no relevance** whether the search engine has **actually knowledge of the fact the personal data are contained in the websites subject to indexing**. Therefore, it amounts to a **data controller** and as such bears the obligations provided by the Directive 95/46



General Data Protection Regulation 2016

Hate speech regulatory attempts?

EU level

Audiovisual Media Service
Directive

Regulation on Terrorism
Content (Proposal)

National level

Germany - Network
Enforcement Act (NetzDG)

Italy – AGCOM Regulation
on hate speech

Audiovisual Media Service Directive

Exemption of liability and appropriate measures – Art. 28(b)(1)

Without prejudice to Articles 12 to 15 of Directive 2000/31/EC, Member States shall ensure that video-sharing platform providers under their jurisdiction take **appropriate measures** to protect:

- (a) minors from programmes [...] which may **impair their physical, mental or moral development** in accordance with Article 6a(1);
- (b) the general public from programmes [...] containing **incitement to violence or hatred directed against a group of persons** or a member of a group based on any of the grounds referred to in Article 21 of the Charter;
- (c) the general public from programmes [...] containing content the dissemination of which constitutes an **activity which is a criminal offence under Union law** (e.g. Article 5 of Directive (EU) 2017/541)

Audiovisual Media Service Directive

No general obligation to monitor – Art. 28(b)(3)

*[...] Those measures shall be practicable and proportionate, taking into account **the size of the video-sharing platform service and the nature of the service that is provided**. Those measures shall not lead to any ex-ante control measures or upload-filtering of content which do not comply with Article 15 of Directive 2000/31/EC. [...]*

BUT

No editorial responsibility (i.e. service providers), but appropriate measures to protect minors and the general public

Regulation on Terrorism Content (Proposal)

A new system without amending the e-Commerce Directive

This Regulation is without prejudice to Directive 2000/31/EC.

Hosting service providers **shall act in a diligent, proportionate and non-discriminatory manner and take into account the fundamental importance of the freedom of expression, the freedom to receive and impart information and ideas in an open and democratic society and with a view to avoiding removal of content which is not terrorist.**

These duties of care shall not amount to a general obligation on hosting service providers to monitor the information they transmit or store, nor to a general duty to actively seek facts or circumstances indicating illegal activity.

Hosting service providers shall remove terrorist content or disable access to it as soon as possible and within one hour from receipt of the removal order from the competent authority.

The German case

NetzDG and hate speech (2017)

- **Aim:** holding social media networks responsible for hosting fake news and hate speech
- **Scope:** Social media networks (with more than 2 million users) must remove hateful content or fake news within a 24 hour time period or face fines. Penalties consisting in fines up to € 50,000,000
- Specifically cites “the experience in the US election campaign” as a reason for a crackdown on “punishable false reports (‘fake news’).”
- **Not comparable with the Russian approach to disinformation**

The French case

Lutte contre la désinformation (2018)

- **Aim:** combating the massive and rapid dissemination of false news via digital tools, including the spread pipes of social networks and media under the influence of a foreign State
- **Scope:** the law addresses specifically those attempts of deliberate information manipulation – targeting the act of dissemination of false information and not the author of said information
- **“action judiciaire en référé”**, emergency proceedings where the judge will qualify the “false news” and order the deletion of the content in question within 48 hours, according to which the contested falsehood must be: (i) obvious; (ii) disseminated massively and artificially; (iii) lead to the disturbance of public peace or the sincerity of an election.

The Italian case – AGCOM Regulation DELIBERA N. 157/19/CONS

AGCOM Regulation

- **Providers of audiovisual and radio video services** subject to Italian jurisdiction are called to ensure respect for **human dignity** and the **principle of non-discrimination** and contrast to instigation to violence and hatred towards groups of people or members of such groups defined with reference to sex, racial or ethnic origin, religion, disability, age or sexual orientation and any other characteristic or personal situation
- The Authority promotes, coordinates and directs the development of codes of conduct by identifying forms of co-regulation, supervision and verification of activities in relation **to video-sharing platform providers**

The Italian case – Co-regulation

AGCOM Regulation

- Audiovisual and radio media service providers are called to guarantee "**the most rigorous respect**" of the fundamental principles enshrined to guarantee users and, especially, human dignity against all forms of discrimination.
- These providers are called to avoid any expression, such as, for example, the titles, which can directly or indirectly "**incite, promote or justify**" the hate speeches the discrimination that offends human dignity or, in extreme cases, can lead to violence, disorder and crime against a person or groups of people.
- The dignity of the individual should be balanced with freedom of expression. The Authority monitors compliance with the provisions of the Regulation in respect of the rights and **freedom of expression and information**.

The Italian case – Co-regulation

AGCOM Regulation

- With regard to **video-sharing platforms**, the AGCOM Regulation does not introduce a catalogue of specific provisions, but provides that the Authority promotes, **through co-regulation procedures, the adoption by these suppliers of measures aimed at combating online distribution**, and in particular on social media, content in violation of the principles established to protect human dignity and for the removal of hate content.
- The obligation to video-sharing platforms consists of:
 - the submission to AGCOM of a **quarterly report on the monitoring carried out for the identification of hate content**, indicating also the procedures and the verification systems in place.
 - The adoption of **effective systems** to identify and notice of unlawful content and wrongdoers.

Conclusions

- Shift from addressing disinformation and hate speech mainly through self-regulation to co-regulation or hard regulation
- Disinformation and hate speech measures adopted by the EU have not amended the e-commerce Directive system unlike the Copyright Directive
- Member States have implemented their strategies to address the phenomena of hate speech and disinformation
- Do we need a comprehensive review of the e-Commerce Directive to avoid legal fragmentation, judicial jeopardization and mitigate the failure of self attempts?
- New challenges for digital constitutionalism
- Proportionality is the key principle