Disinformation, Fake News and Deep Fakes.

Special Raporteur, Group II, AIDP

Tommaso Guerini
Pegaso Telematic University of Naples
tommaso.guerini@unipegaso.it

1. Introduction

Korean philosopher Byung-Chul Han observes in a recent study how "*the current hyperinflation of objects, leading to their explosive proliferation, is in turn a symptom of a growing indifference to them. Our obsessions are no longer directed at things, but rather at information and data. We now produce and consume more information than things. We become inebriated with communication. Libidinal energies abandon things and throw themselves on the not-things. The consequence of this is called* infomania."[1] .

This is an extremely useful key to introduce the topic of the relationship between massive disinformation carried out through digital tools and criminal law.

In a time seemingly close but now terribly distant, before a pair of *black swans*[2] appeared on the horizon, irreversibly undermining the fragile certainties that had characterized *homo occidentalis* in the decades following the *end of history*[3] and making the current geopolitical framework *plutôt floue*[4] , it was necessary to pose the problem of identifying an autonomous semantic space for the phenomenon we are about to investigate, with respect to the millennial tradition of 'common fake news' and its impact on the public sphere.

[1] B-C. HAN, *Le non cose*, Torino, 2022, 7.
[2] A term that, beginning with N.N. TALEB's essay, *The Black Swan,* Milan, 2014, designates a rare event with a huge impact and predictable retrospectively, but not prospectively, we refer specifically to the *Covid-19* pandemic that developed starting in early 2020 and Russia's invasion of Ukraine that began on Feb. 24, 2022, and is still ongoing at the time we dismiss this writing.
[3] The expression, as is well known, was coined after the fall of the Berlin Wall by F. FUKUYAMA, *The End of History and the Last Man*, New York, 1992; the author has recently returned to some of his theories, revising them in part, in ID., *Identity: The Demand for Dignity and the Politics of Resentement*, New York, 2018.
[4] R. QUENEAU, *Les Fleures bleues*, Paris, 1975, 3.

Even before the digital revolution and the advent of what the Italian philosopher Luciano Floridi has called the *Infosphere*[5] , which as we shall see represent the background of the phenomenon we are observing, manipulation – of individuals, of the masses – has always been an extremely effective tool of political struggle, which both authoritarian regimes and democracies – authoritarian and otherwise – have never really given up, and its evolution goes hand in hand with that of both political and social systems and the technologies available to amplify its effects.

Since the dawn of time[6] , those wielding power have been concerned with the control of the masses: from ancestral religions to contemporary times, *elites* have worked out gradually more refined systems to obtain and defend their supremacy.

The ever-increasing prominence of the popular masses in politics since the Liberal Revolutions of the 17th Century[7] meant that both democracies and authoritarian regimes worked to develop increasingly effective *propaganda* tools.

While, as Edward L. Bernays observed, "*the conscious and intelligent manipulation of the opinions and habits of the masses plays an important role in a democratic society*" so much so that "*those who master this social device constitute an invisible power that truly runs the* country,"[8] , on the other hand, as Hannah Arendt observed, propaganda is an indispensable tool for totalitarian regimes[9] .

We are thus faced with a *constant*, the conformation of which changes as the sociopolitical context of reference changes.

Contributing to making the present profoundly different from even the most recent past is the interplay between two recently arising phenomena.

On the one hand, on the economic even before the political level, the Western model of democracy is being openly challenged by the so-called *illiberal democracies*[10] , above all Russia and China, but also, on our continent, by Poland and Hungary and, alas, to some extent today also by Italy.; on the other hand, the field of confrontation between the different political-economic models has already shifted from the world of atoms to

---

[5] L. FLORIDI, *La quarta rivoluzione. Come l'infosfera sta trasformando il mondo*, Milano, 2017; ID., *Pensare l'infosfera. La filosofia come design concettuale*, Milano, 2020.
[6] We refer for all to E. CANETTI's rich essay, *Mass and Power*, Milan 1960-2011.
[7] According to Hannah Arendt: "*totalitarian propaganda perfects the technique of mass propaganda, but does not invent or propose its themes. These are already ready, prepared by the fifty years of the rise of imperialism, the disintegration of the nation state and the appearance of the plebs on the European political scene.*" H. ARENDT, *The Origins of Totalitarianism*, Turin, 2009, 484
[8] E.L. BERNAYS, *Propaganda,* New York, 1928, 25.
[9] H. ARENDT, *The Origins of Totalitarianism,* cit., 471 ff.
[10] F. FUKUYAMA, *Identity,* Milan, 2018, 19.

the world of the *bit* - with particular attention to the issues of *Artificial Intelligence*[11] - and the trend, also due to the strengthening of infrastructures linked to the migration of an increasing number of activities from the analog to the digital universe, which occurred during the Covid-19 pandemic, now seems unstoppable, making the role of social networks increasingly central in every sphere of human life[12] .

The power of the image of Jake the Shaman breaking into the sacred precinct on *Capitol Hill* on January 6, 2020 lies in mixing the apotropaic power of primitive man-the gestures, the headgear, the tattoos-with the esoteric power of the digital dimension.

Just as in *primitive* cultures, the shaman represented a bridge between the dimension of the divine and the earthly, so in the age of the *infosphere*[13] constitutes a conduit for interacting with the digital universe, attributing substance to what would otherwise be destined to remain in the world of *bits*.

It is no coincidence that in the United States Jake Angeli is called *Q Shaman*[14] , because of his conspiracy theory propaganda activities named *Qanon*.

Beyond the folkloric aspects of the affair and the little impact it had on the regularity of the electoral process - having in fact only caused a delay of a few hours in the proclamation of the elected president - the impact of the syncretism between real and virtual underlying it should not be underestimated.

Qanon was born on the Internet and propagated through *social networks*, thanks in part to the incessant support - more or less overt - of the President of the United States, who on numerous occasions relaunched with his *tweets* the theses elaborated by *Q*'s followers, going so far as to provide them - with the systematic dissemination of real *fake news*[15] on the conduct of the November election process - with additional elements to

---

[11] As most recently argued by KAI-FU LEE, *AI Superpowers. China, Silicon Valley and the new world order,* Boston, 2018, the frontiers of the digital economy-and in particular the potential offered by the development of *Artificial Intelligence*, or AI-represents the new battleground between the superpowers of the 21st century and, in particular, the United States and China.

[12] For some considerations on the role of *social networks* in the Covid-19 pandemic and the possible effects of the migration into the digital universe of activities traditionally carried out "in presence," be granted the reference to T. GUERINI, *La tutela penale della libertà di manifestazione del pensiero nell'epoca delle fake news e delle infodemie*, in *this Review*, 2-2020, 205-207.

[13] In this section and in the remainder of the text we will refer several times to the thought of philosopher Luciano Floridi, who has observed how the constant interaction between *online* and *offline* dimensions is producing an irreversible transformation of the conceptual architectures of the present. L. FLORIDI, *The Fourth Revolution. How the infosphere is transforming the world,* Milan, 2017, 99-105; ID., *Thinking the infosphere. Philosophy as conceptual design*, Milan, 2020.

[14] W. WRIGHT, *Trump Loyalists Arrested and Charged With Involvement in Capitol Siege*, in *The New York Times*, January 9, 2021.

[15] It seems to us, in fact, that in the case of the *Stolen Election* thesis, all the elements that characterize *fake news* recur, which in our perspective are characterized by being information that is partly or wholly untrue, produced and disseminated intentionally or unintentionally through the Web, media or digital

support the fact that there are in the public administration occult forces engaged in criminal activities aimed at undermining the democratic order.

Just as with *fake news*, this *conspiracy theory, which is* self-feeding in a digital ecosystem and then overflowing into the *real* world, can also be seen as an *original* and *originating* product of the digital revolution.

Thus, some *topoi of* contemporary thought come to the fore, with which the jurist is also required to deal.

The digital revolution-which, again according to Floridi, represents the "fourth revolution" in the history of thought[16] -has brought about, to quote Antoine Garapon and Jean Lassegue's effective image, an *anthropological rupture*[17] , which is characterized by the progressive ability of the *online* digital world to overflow into the *offline* analog world, with which it mixes giving rise to the *onlife* experience; a *tertium genus* in which the very distinction between analog and digital ends up losing its meaning[18] .

The progressive development of ICT means that more and more people in the world are living in a new dimension-the *infosphere-which is* increasingly *synchronized*, *delocalized* and *correlated*[19] .

As noted more recently by the author himself, in this context the problem of rules-which was guiltily ignored when the decision was made in the 1990s not to introduce limits to the Internet so as not to harness its growth-becomes central[20] .

---

communication technologies, and characterized by an apparent plausibility, the latter fueled by a distorted system of public opinion expectations and an amplification of underlying prejudices, which facilitates their sharing and dissemination even in the absence of source verification, such as to harm individual legal assets, such as honor and reputation, or aimed at affecting, directly or indirectly, the freedom of citizens to exercise their right to vote and to affect the proper functioning of democratic institutions. T. GUERINI, *Fake news and criminal law,* cit., 28-29.

[16] The first *revolution came about* thanks to the thought of Nicolaus Copernicus, through which an initial, radical rethinking of man's role in the Universe was brought about. It was followed by a *second revolution*, which took place starting in 1859, the year of publication of Charles Darwin's *The Origin of Species*, in which it was shown that every living being was the result of evolutionary dynamics that are realized through a process of natural selection, thus removing the human being from the position of the barycenter of the biological realm. Finally, the *third revolution*, is that due to the thought of Sigmund Freud, who, through the development of psychoanalytic thought, challenged man's view of himself. Instead, the thought and work of British mathematician Alan Turing (and, in particular, the essay *On Computable Numbers, with an Application to the Entscheidungsproblem*, in *Proceedings of the London Mathematical Society*, Volume s2-42, Issue 1, 1937) would give rise to the *fourth revolution*, deposing man from the privileged and exclusive position he held "*in the realm of logical reasoning, the ability to process information and to act intelligently*. L. FLORIDI, *The Fourth Revolution. How the infosphere is transforming the world,* Milan, 2017, 99-105

[17] We borrow the effective image coined by A. GARAPON-J. LASSÈGUE, *Digital Justice,* Paris, 2018, 16 ff.

[18] L. FLORIDI, *The Fourth Revolution*, cit., 47-48.

[19] *Ibid*, 53.

[20] L. FLORIDI, *The Green and the Blue. Naive ideas for improving politics,* Milan, 2020, 67.

It is in this context that *fake news*, *deepfakes* and, in general, digital tools of massive disinformation arise and proliferate, which in our study perspective are nothing more than a particularly insidious evolution of traditional political propaganda tools.

Fake news and *deepfakes* are technologically evolved consensus manipulation tools to cope with the needs that have arisen with the digital revolution, in which it is no longer enough to direct the *analog* lives of citizens, but rather to condition their behavior in the Infosphere.

A further order of considerations may be useful to us in directing more precisely the more properly "criminalistic" considerations that we propose to carry out in the coming chapters.

Both the chronicles of the pandemic and the most recent images emanating from the Ukrainian front reinforce the thesis that through the spread of *fake news,* legal goods, both individual and collective, of primary importance can be harmed, including life, public health, and even the proper functioning of the constitutional organs of the state.

With reference to the field of public health, it is now well known how, since the first news about the virulence of Covid-19 spread, the *World Health Organization* (WHO), in parallel highlighted the dangers arising from what has been called the first *infodemic* in human history[21] , consisting of the massive dissemination of *fake news* about the virus, such that it gave rise to a *digital swarm*[22] suitable for concealing the news that government *agencies-primarily*, health authorities-needed instead to disseminate in order to control the epidemic and mitigate its impact on the population[23] .

On the other hand, as for the Russian-Ukrainian war, every day there are news reports about the spread of *fake news* and *deepfake*[24] , aimed at destabilizing the Ukrainian

---

[21] J. ZAROCOSTAS, *How to fight an infodemic*, The Lancet, Vol. 395, Feb. 29, 2020, 676. The neologism describes, according to its definition in the Treccani Vocabulary, the phenomenon of "the *circulation of an excessive amount of information, sometimes not sifted accurately, that makes it difficult to find one's way around a given topic because of the difficulty of identifying reliable sources.*" See *Infodemia* (entry), in *Vocabolario on line*, at www.treccani.it/vocabolario. The lemma originated in English (*infodemic*) as a word composed of the nouns Information and Epidemic. It would be an *author's word*, coined by David J. Rothkopf, a Washington Post journalist, in a 2003 article titled *When the Buzz Bites Back*, which was about the 2002-2004 SARS epidemic. See www1.udel.edu/globalagenda/2004/student/readings/infodemic.html

[22] B-C. HAN, *In the Swarm. Visions of the digital,* Rome, 2014.

[23] We have dealt with the topic in our: T. GUERINI, *La tutela penale della libertà di manifestazione del pensiero nell'epoca delle fake news e delle infodemie*, in *DisCrimen*, June 15, 2020, to which we refer for further discussion.

[24] The neologism, coined in 2017, crosses the locution *deep* learning (set of techniques that allow Artificial Intelligence to learn to recognize shapes) with the word *fake* (false, fake news) - and describes those techniques that allow images to be combined and superimposed on each other using generative neural

population and army[25] , i.e., manipulating Western public opinion in favor of one side or the other[26] .

Indeed, the recognition of the war potential of *fake news* is not a feature of contemporary times, if anything, an element of affinity with the historical tradition of fake news.

As Sun Tzu warned, "*War is the Tao of deception. Therefore, if you are skillful, in the face of the enemy feign inability. If you are forced to commit your forces, feign inactivity.If your target is close, make believe that he is far away; when he is distant, create the illusion that he is* nearby. "[27] .

We must therefore take note, on the one hand, that human history has been and is constantly dotted with events determined by the dissemination and use of fabricated information[28] , and likewise, on the other hand, recognize how, within this category, *fake news* represents a new *paradigm*, in the Kuhnian sense[29] .

2. Disinformation, *fake news* and *deep fakes.*

Having acknowledged the originality of the phenomenon and reconstructed what constitute its essential elements, in order to take our study a step further, it is necessary to make a categorization of *fake news in* order to assess its possible intersections with current criminal law, that is, what goods or interests are affected by their dissemination that lack a criminal law safeguard, and this in order to assess the compatibility with our constitutional set-up of any *new* incriminations.

networks (in technical jargon, GANs). Cf. N. ORDONSELLI, *"Porn Deepfake": criminal law profiles*, in *Cyberlaws*, January 18, 2021.

[25] See, for example, the *deep fake* video in which Ukrainian President Volodymyr Zelensky ordered his country's armed forces to lay down their weapons in front of the Russian military. See: https://www.youtube.com/watch?v=X17yrEV5sl4

[26] We realize that this statement requires clarification. As we shall see at length *infra,* § X, the Russian government has for many years been the subject of attention by international agencies designed to monitor respect for freedom of opinion and information. The most obvious form of linguistic manipulation is on the Russian side's choice to deny the very existence of a *war*, calling the invasion of Ukraine a "*special military operation.*" However, it is safe to assume that there is also on the Ukrainian side the use of digital disinformation tools aimed at reinforcing the repeated calls for relief from the international community.

[27] SUN TZU - SUN PIN*, The Art of War*, Vicenza, 2005, 127.

[28] In a valuable volume published a few years before the syntagma *fake news* entered common parlance, Luciano Canfora offered an insight ranging from Pausanias' letter to Xerxes to the "strange" letters sent in 1928 by Ruggero Grieco to Umberto Terracini and Antonio Gramsci during the trial to which they were subjected by the fascist regime. L. CANFORA, *La storia falsa,* Milan, 2008.

[29] T. S. KUHN, *The structure of scientific revolutions*, Turin, 1978.

Although we are moving in a particularly uncertain terrain-or rather, precisely because of the fact that the phenomenon we are discussing presents itself as difficult to harness in a satisfactory definition from a scientific point of view-we believe that for a proper understanding of the physiognomy of the current information ecosystem it is necessary, before examining the motivations underlying their production and dissemination and the ways in which this occurs, to outline a possible taxonomy of the main forms of manifestation of the phenomenon.

According to a model developed in the United States by Claire Wardle, one of the most influential scholars of *user generated content*[30] , seven different types of news stories can be distinguished as having the character of *disinformation* or *misinformation*[31] , which are distinguished in that in the former case there is a misleading intent, which is absent in the latter[32] .

At the base of the pyramid, we find fake news spread with a satirical *or parody* purpose (*satire or parody*), that is, characterized by having no offensive purpose and being potentially amusing.

Then there is the case of dissemination of *misleading content, which* is when information-true or likely-is manipulated or distorted in order to simplify complex concepts, altering their exact perception by end users.

The existence of a more or less extensive core of truth is also a common feature in cases of fraudulent content (*imposter content*), in which the alteration concerns the source of the news; false connections (*false connection*), which occur whenever the headlines or visual content accompanying the article do not correspond to the content of the article; and, finally, cases of false context (*false context*), which occur when other

[30] User-generated content is defined as any textual or visual content of a creative nature generated directly by a user in a public platform, such as, for example, *Facebook*, *Twitter* or *YouTube.* As has been noted, UGC has rapidly revolutionized the way most people relate to traditional media - above all, television - reshaping the Web 2.0 landscape. For an in-depth discussion, see J. VAN DIJCK, *Users like you? Theorizing agency in user-generated content*, in *Media, Culture & Society*, 2009, Vol. 31(1), 41 ff.; T. DAUGHERTY-M.S. EASTIN-L. BRIGHT, *Exploring consumer motivations for creating user-generated content*, in *Journal of Interactive Advertising,* Vol. 8, No. 2, 2008, 16 ff.

[31] C. WARDLE, *Fake news. It's complicated,* in *www. firstdraftnews.org*, February 16, 2017.

[32] In other words, while *misinformation* consists of the accidental dissemination of false news, *disinformation is* characterized by the voluntariness of the conduct. See on this point: B.C. STAHL, *On the Difference or Equality of Information, Misinformation, and Disinformation: A Critical Research Perspective,* in *Informing Science: The International Journal of an Emerging Transdiscipline*, Vol. 9, 2006, 83 ff. As we will elaborate further in the remainder of our discussion, in the mechanism of operation of *fake news* the two notions may present a complementary character, as both *conscious* and *unconscious* diffusers can interact.

false information, even if of lesser importance, is disseminated to accompany a true news story.

There are two cases that we can define as extreme, found in the case where an information content, image, or film is expressly manipulated in order to deceive the public (*manipulated content),* or in the case of fabrication of false news (*fabricated content*), which occurs in all those circumstances where an information, image, or film is completely false and intentionally made to deceive the users.

From a criminal law perspective, however, it seems to us necessary to examine the problem from a different angle, which must necessarily take into consideration both the object of the content and the purpose by reason of which it is formed and shared.

Reasoning in terms of offensiveness, in fact, we can identify fake content - thus belonging to all seven of the different types of *fake news* examined so far - that is completely incapable of harming goods or interests of a constitutional nature, as well as seemingly neutral *fake news that is* nevertheless capable of endangering or damaging legal goods of an individual and super-individual nature.

In our study perspective, we will therefore distinguish: (a) content of a *defamatory* nature, such being all cases in which a piece of news or visual content harms the honor or reputation of one or more identified or identifiable individuals, without the justification grounds of the right to chronicle and the right to criticism finding application; (b) *revisionist-type* historical content[33] , i.e., false information disseminated for the purpose of influencing the thinking of contemporaries regarding events of the past[34] ; c) false content of *an anti-scientific* nature - i.e., aimed at supporting minority theses or theses not accepted *tout court* by the specialized literature, not epistologically verifiable and falsifiable[35] ; d) generalist or entertainment *fake news*, *apparently* devoid of an offensive

[33] As we have already made clear (see *supra,* § 5) we have decided to exclude from this study the examination of the criminal profiles related to the phenomenon of *denialism*. We therefore adopt the term *revisionist*, to indicate those false, or at any rate manipulated, news items which, while not ascribable to - legitimate - historiographical criticism, do not, by reason of their content, rise to *denialism* proper. For a precise examination of the semantic issues underlying the problem of denialism, see E. FRONZA, *Il negazionismo come reato*, cit., XIII-XVI.

[34] A typical example of this category is the *fake news* according to which it was Benito Mussolini who "gave the Italians a pension" by inventing the welfare system. see instead F. FILIPPI, *Mussolini also did good things*, Turin, 2019, 7 ff., where he reconstructs the history of the Italian welfare system, established in 1895 by the Crispi government. Lastly, he dealt with the theme of truth in the historical field - tracing the events of thirty historical episodes subject to *manipulation* - P. MIELI, *Le verità nascoste,* Milano, 2019.

[35] See on this point, among others, the study by S. DI GRAZIA, *Medicines and Lies*, Rome, 2017.

content[36] and disseminated in order to generate traffic on the sites that carry them, thus realizing profits through the mechanism of the c.d. *clickbaiting*[37] .

Indeed, it seems to us that this hypothetical taxonomy may help to capture some offensive nuances that more concise lists risk leaving out[38] .

In view of the above, by the term *fake news* in the criminal context we refer to: "*a piece of information that is partly or wholly untrue, produced through the use of digital data and disseminated intentionally or unintentionally through the Web, media or digital communication technologies, and characterized by an apparent plausibility, the latter being fueled by a distorted system of public opinion expectations and an amplification of the underlying prejudices facilitating its sharing and dissemination even in the absence of source verification*, *such as to endanger or injure individual legal assets, such as honor and reputation, physical integrity and life*, *or more frequently aimed at affecting, directly or indirectly, the freedom of citizens to exercise their right to vote and to affect the proper functioning of* democratic *institutions.*"[39] .

It seems to us that this definitional proposal succeeds in holding together *contextual elements*, such as the sociocultural and technological formant, and *ontological* elements or *structure* of *fake news*, which cannot be 'really' considered news, lacking both truth requirement and the qualification of facts pre-existing the news activity.

A peculiar form of deceptive content is *deep fakes*, which we have alluded to several times in the preceding pages.

The term *deepfake* originated from the crasis between deeplearning (referring to the learning ability of an Artificial Intelligence) and *fake*[40] .

---

[36] In contrast, an in-depth study conducted by R. MUIRHEAD-N.L. ROSENBUM, *A lot of people are saying*, Princeton, 2019, demonstrates that there is no such thing as fake *news that is* truly devoid of offensive content, as the dissemination of falsehoods that are not directly damaging to individual or collective goods or interests nonetheless results in lowering users' defenses, preparing them to receive other - and more dangerous - falsehoods.

[37] For an analytical review of the main *hoaxes* spread through the net, see the paper by D. PUENTE, *The Great Internet Deception*, cit.

[38] Consider, for example, the distinction proposed by M. BASSINI-G.E. VIGEVANI, *First notes on* fake news *and its surroundings*, in *Media Laws*, 1-2017, 16, where the two Authors distinguish between three categories of content. The first includes falsehoods artfully constructed by power groups, sometimes by foreign governments, with the aim of changing the public agenda; the second includes fake news or news of dubious authenticity that circulates on the web supported by sharing among users (the *vox populi*); and the third includes all fake news that harms individual or collective interests.

[39] This is a necessary update, in light of what was argued *supra*, of the definition we had come up with in 2020. Cf. ibid. 28-29.

[40] M. BEARD, *To fix the problem of deepfakes we must treat the cause, not the symptoms*, in www.theguardian.com, July 23, 2019.

Developed as part of the *online* pornography industry[41] and quickly spreading to the entire *World Wide Web*, through the use of particularly precise algorithms, *deep faking makes* it possible to replace the faces of two people (so-called *face swapping*), producing videos in which anyone can be made to say anything, even lip-syncing perfectly.

The offensive potential of this type of tool is considerably higher than that of any other bogus information product, as the alteration of an audiovisual document has an extremely high deceptive attitude, placing the user in front of something that he sees happening before his eyes, without having any tools to refute its veracity.

Over the past few months, incidents of *deepfake* spreading have been multiplying and range from cinematography[42] to politics.

It is clear that in the latter case the dangers associated with the use of this technology are immediately appreciable, especially in the political-electoral arena[43] .

In fact, the risks to the proper allocation of consensus in the political-electoral sphere seem to us to be particularly concerted, both because of the continuous refinement of the programs that allow for the creation of this content - now technically very accurate, except for some imperfections in the dubbing, which are difficult to perceive by a user who is not particularly attentive - and because of the scarcely repressive policy put in place by the companies that manage the main sites for the dissemination of video[44] .

Exemplary of this attitude of underestimation of the risks for the alteration of the democratic process underlying the dissemination of this technology is the affair involving the *Speaker* of the *United States House of Representatives,* Nancy Pelosi, who through

---

[41] On the role of pornography as a *hidden engine* of the tech industry, see the article by R. BENES, *PORN: The Hidden Engine That Drives Innovation In Tech*, in www.businessinsider.com, July 5, 2013.

[42] For example, a video has been much publicized in which Jack Nicholson's face has been superimposed on Jim Carrey's, pretending that the latter filmed some scenes from Stanley Kubrick's masterpiece, *The Shining* (1980) (See https://www.youtube.com/watch?v=HG_NZpkttXE). Similarly, it is possible to visualize a short clip from Martin Scorsese's film, Taxi Driver (1976), in which the original protagonist-Robert De Niro-was superimposed on the face of Al Pacino (See www.youtube.com).

[43] R. CHESNEY-D.K CITRON, *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*, in *California Law Review* (July 14, 2018). 107 (2019, Forthcoming); U of Texas Law, Public Law Research Paper No. 692; U of Maryland Legal Studies Research Paper No. 2018-21.

[44] In a statement released following the sharing of *deepfake* videos via the *Facebook* platform, the *social network*'s managers argued against the removal of such content, claiming that it is important to the Company that each person can decide what to believe in, while their job is to ensure that users can receive the right information ("*We think it's important for people to make their own informed choice for what to believe. Our job is to make sure we are getting them accurate information.*"). See A. HORTON, *Facebook defends decision to leave up fake Pelosi video and says users should make up their own minds*, in www.washingtonpost.com, May 25, 2019.

the manipulation of some videos was made to appear as if she were speaking in a drunken state[45] .

One such *deepfake* video was shared via *Twitter* by President Donald Trump[46] , garnering over three and a half million views[47] and tens of thousands of subsequent shares, further demonstrating how in cases such as these a central role is played by confirmation bias -- in this case, the conservative electorate's biases against Pelosi, who has long been a major target of *Great Old Party* propaganda.

Having clarified this first and fundamental point, it remains to be ascertained for what reasons and to what extent this phenomenon may be of interest to the criminal law scholar.

A first answer can be given from an *internal* perspective *within* the criminal justice system: in the era that Filippo Sgubbi has called *total* criminal law[48] there are not and cannot be any relevant social phenomena that do not have an impact on criminal law, in comparison with which "*the factual has* normative *impact.*"[49] .

For this reason, even in jurisdictions, which do not expressly provide for a penalty discipline dedicated to the spread of *fake news, it* is not possible to help but wonder how this phenomenon can also affect those who deal with our subject.

There is a second order of reasons, however, makes the use of the comparative formant indispensable in our view.

As we are about to see, numerous jurisdictions have chosen to regulate the phenomenon, including through the use of the punitive instrument.

Understanding how this happened and what legal goods were sacrificed to-attempt to-do so will provide us with a useful tool for thinking about our near future.

3. *Comparative models of fake news regulation between protection of freedom of thought manifestation and censorship thrusts*

[45] In this case, the result was achieved through the manipulation of a real speech by the *Speaker*, whose words - slowed down and resynchronized with the video - produced the impression of a person in a drunken state. See *www.youtube.com/watch?v=sDOo5nDJwgA*.
[46] See *www.youtube.com/watch?v=EfREntgxmDs*
[47] S. COLE, *Americans Don't Need Deepfakes to Believe Lies About Nancy Pelosi*, in www.vice.com, May 24, 2019.
[48] F. SGUBBI, *The total criminal law*, Bologna, 2019.
[49] *Ibid*, 29.

We have had the opportunity to observe how the phenomenon we are discussing is transnational and, without emphasis, global in nature.

Therefore, a reconstruction of the main models adopted to counter the massive dissemination of fake news through the Web is a particularly useful tool for the domestic jurist, who is confronted, on the other hand, with a general backwardness in the legal regulation of technological phenomena.

Before making such an examination, a note of method is required. One of the essential elements of the phenomenon we are investigating is represented by the impossibility of harnessing it within physical boundaries, a feature common to many phenomena developing in the *World Wide Web* ecosystem. The problem of the absence of territorial ties within the Internet, which is in some ways a *classic* theme with respect to the needs for legal regulation of behavior enacted in or through the network[50] , takes on, in the context of *disinformation* studies, an entirely peculiar dimension. Just as the network has no physical boundaries, neither do *fake news, the* spread of which has taken root - albeit with some differences - in all countries in which there are sufficiently developed information infrastructures.

Remaining in the familiar sphere of criminal law, it is now well established that there are criminal phenomena that are transnational in nature, such as smuggling, drug trafficking or money laundering, which are now widely studied and recognized in their essential elements, both criminologically and from the point of view of legal recognition through international Conventions.

Conversely, beyond a generic acknowledgement of the danger to the proper functioning of the information system caused by *disinformation* campaigns, the

[50] We will, moreover, have the opportunity to spend some considerations on the progressive formation - in spite of certain *cyberlibertarian* orientations that would like the network to be open and universal - of a plurality of networks within the Internet, divided by linguistic, social and political barriers. The formation of different networks in the network is one of the phenomena of greatest interest at this historical moment, as it seems to re-propose - updating them - some typical stylistic features of the confrontation between superpowers that marked the second half of the twentieth century. (See on this point F. FUKUYAMA, *The End of History and the Last Man,* New York, 1992 (transl., *La fine della storia e l'ultimo uomo*, Milan, 2003); the author has recently returned to some of his theories, revising them in part, in ID., *Identity: The Demand for Dignity and the Politics of Resentement*, New York, 2018 (transl., *Identity. The Quest for Dignity and the New Populisms*, Milan, 2019). As most recently argued by KAI-FU LEE, *AI Superpowers. China, Silicon Valley and the new world order*, Boston, 2018, the frontiers of the digital economy - and in particular the potential offered by the development of *Artificial Intelligence*, or AI - represent the new battleground between the superpowers of the 21st century and, in particular, the United States and China.

phenomenon of information distortion still remains largely unexplored and the subject of point-in-time interventions, entrusted to the sensitivity of individual jurisdictions.

In the following pages we will acknowledge the legislation adopted in three different geo-political quadrants: Europe, Russia and Asia, and this is because of the gradual delineation in the *magnum sea* of *cyberspace of* material and immaterial boundaries, which delineate a contrast between eastern and western networks.

Belonging to the first category are the technological barriers that allow individual governments to restrict users' access to certain sites or access to certain services[51] , due to which today we can observe, alongside the primal model of a Western-driven Internet (U.S. and European Union), the gradual emergence of a Russian and a Chinese network[52] .

Instead, we define *intangible* barriers as those pertaining to economic, linguistic and cultural dimensions, such as to differentiate the approach of users and the dynamics of their interactions through the network. The history of the Internet was written in the English language[53] , but already for several decades the main portals, sites and *social networks have* been available in most of the languages spoken around the globe; however, this does not prevent certain languages from acting in themselves as barriers to the proliferation of *fake news* in a given country, as evidenced by the fact that in Japan - according to a study conducted by the *Law Library of Congress* - the phenomenon of spreading fake news has remained substantially limited to the traditional *media*, partly due to the fact that: "*the difficulty of the Japanese language for foreigners prevents the phenomenon of* fake news *publication by people located outside* Japan."[54] .

3.1.- *Freedom of Thought and* Disinformation *in the European Conventional Framework*

[51] The policy of restricting services provided by Western *corporations* has long been adopted by the Chinese government-whose sovereignty spans more than one-sixth of the global population-which, for example, does not make many applications developed by *Google* accessible within the territory of the People's Republic of China, including the *Gmail* e-mail service, *Google maps,* or the machine translation application *Google Translate.*

[52] J. NOCETTI, *Contest and conquest: Russia and global internet governance*, in *International Affairs*, Vol. 91, 1-2015, 111 ff.

[53] J. NAUGHTON, *A Brief History of the Future. The origins of the internet*, London, 1999.

[54] The Law Library of Congress, *Initatives to Counter Fake News in Selected Countries*, April 2019, 52.

In the European Convention on Human Rights (ECHR), the right to free expression of thought is enshrined in Art. 10, where every person is recognized as having the right to freedom of expression, but which at the same time dictates the requirements that any rules providing for possible forms of limitation to this freedom must meet, identified in the reservation of law, *proportionality* and reservation of subject matter[55] ; exceptions which, already in the earliest case law, have been interpreted by the European Court of Human Rights in an extremely restrictive key, recognizing freedom of expression as having a wide margin of effectiveness.

For the Court, since its first ruling dealing with Article 10, § 2 of the Convention[56] , freedom of expression is a bulwark of democratic society, as well as one of the fundamental elements for its development and for the progress of every human being, while with reference to freedom of the press, defined as one of the best tools available to citizens to know and judge the ideas and actions of those who govern them, it observed that "*freedom of political debate is at the heart of the concept of democratic society that pervades the entire* Convention."[57] .

We should note, however, that the concerns expressed by the Court regarding the freedom of manifestation of thought, as enshrined in Article 10 of the Convention, seem to be less present in more recent case law, which has dealt specifically with the exercise of that freedom in the context of the Internet.

The problem of the increased offensive capacity of the new media and information - in part already manifested in the *Jersild v. Denmark* judgment[58] - underlies a *restrictive* view of the scope of freedom of expression in this specific area, as can already be seen in

---

[55] "*Everyone has the right to freedom of expression. This right includes freedom of opinion and freedom to receive or impart information or ideas without interference by public authorities and without frontier limits. This article does not prevent states from subjecting broadcasting, film or television enterprises to a licensing regime.*" According to § 2: "*The exercise of these freedoms, since it entails duties and responsibilities, may be subject to such formalities, conditions, restrictions or sanctions as are prescribed by law and constitute measures necessary, in a democratic society, for national security, territorial integrity or public safety, for the defense of order and the prevention of crime, for the protection of health or morals, for the protection of the reputation or rights of others, to prevent the disclosure of confidential information, or to ensure the authority and impartiality of the judiciary.*"

[56] EDU Court, Handysyde v. United Kingdom, December 7, 1976, § 49.

[57] EDU Court, Lingens v. Austria judgment, July 8, 1986, § 42.

[58] EDU Court, *Jersild v. Denmark* judgment, Sept. 23, 1994. In which the Court had previously observed that "In considering the '*duties and responsibilities' of a journalist, the potential impact of the medium in question is an important factor and it is commonly recognized that audiovisual media often have a much more immediate and powerful effect than print media.*" EDU Court, *Jersild v. Denmark* judgment, cited above, § 31.

the *Editorial Board of Pravoye Delo and Shtekel v. Ukraine, in which* there is a manifest concern about the risks of harm caused by Internet content and communications to the exercise and enjoyment of human rights and freedoms, particularly the right to respect for private life, which are undoubtedly greater than those posed by the press[59] .

This same fear is also read against the grain in a controversial[60] ruling (*Delfi v. Estonia*), in which the ECHR noted how the Internet, while constituting an unprecedented platform for the exercise of freedom of expression, can also lead to the emergence of new dangers, such as the dissemination in ways once unthinkable, worldwide, in a matter of seconds, of defamatory or otherwise unlawful content, including speech inciting hatred or violence, which moreover may remain persistently available *online*[61] .

In order to complete the picture of Strasbourg's jurisprudence on freedom of the press, it is necessary to take into account a final group of pronouncements, in which the ECHR has dealt with the legitimacy of prison sentences for cases of journalistic defamation, an issue that is intertwined with some cases of great interest to the domestic jurist (see *below*, § 4.1.) and with some considerations that we will set out in our conclusions (see *below*, Ch. IV, § 4.2.).

This is primarily the *Cumpănă and Mazăre v. Romania*, a true *leading case on the* proportionality of the sanction against press freedom cases, in which the EDU Court held that "*although States Parties are authorized, or even obliged, by their positive obligations under Article 8 of the Convention (...) to regulate the exercise of freedom of expression in such a way as to ensure adequate legal protection for the reputation of individuals, they must not do so in such a way as to unduly deter the media from fulfilling their role*

[59] EDU Court, Judgment *Editorial Board of Pravoye Delo and Shtekel v. Ukraine*, May 5, 2011, § 63.

[60] O. POLLICINO, *The constitutional perspective on freedom of expression in the internet age*, in G. Pitruzzella-O. Pollicino-S. Quintarelli, *Words and power,* Milan, 2017, 12-13.

[61] EDU Court, *Delfi v. Estonia* judgment, June 16, 2015, § 110. The judgment is published in *Quad. cost.*, 2014, 457-459, with a note by G. E. VIGEVANI, *The civil liability of sites for anonymous writings: the case Delfi v. Estonia*. The principles of *Delfi v. Estonia* were subsequently taken up-although with some distinctions that are not always immediately understandable-by the subsequent jurisprudence of the EDU Court, in particular in MTE *v. Hungary* (EDU Court, *MTE v. Hungary*, Feb. 2, 2016, in *Quad. cost*, 2016, 393-400, with a footnote by S. VIMERCATI, *Magyar v. Hungary: the European Court returns to the liability of* web *portals*) and and *Rolf Anders Daniel Pihl v. Sweden* (EDU Court, *Rolf Anders Daniel Pihl v. Sweden* judgment, March 9, 2017, for a commentary on which: S. VIMERCATI, *The Strasbourg Court returns to site operator liability: the case of Rolf Anders Daniel Pihl v. Sweden*, in *MediaLaws - Rivista dir. Media,* 1-2017, 152 ff.) in which, most recently, it was reiterated that the manager of a *blog* can be held liable for the publication of a defamatory comment entered by a user who has remained anonymous, if it does not contain expressions that transmogrify into incitement to hatred and violence and has promptly, following the report of the person concerned, removed it.

of informing the public about obvious or suspected cases of misuse of public *power."*[62] . According to the Court, in such cases, the *chilling effect* that the fear of such severe sanctions brings about on the exercise of freedom of expression in journalism is evident and that, therefore, recourse to the deprivation of freedom sanction is permissible only in the presence of exceptional circumstances, exemplified by the cases of *hate speech or incitement to violence* and not, on the other hand, in the case of *common* defamation in the press.

Equally significant of the Strasbourg Court's restrictive orientation with respect to the application of prison sentences to journalists are two additional judgments: *Fatullayev v. Azerbaijian*[63] , in which a violation of Art. 10 ECHR for the prison sentence imposed on two journalists, again for failure to comply with the requirement of *proportionality*, and *Katrami v. Greece*[64] , in which the violation of the Convention was enshrined with reference to a suspended prison sentence, which was nevertheless considered disproportionate by the Court.

Thus, a kind of *double track is* emerging: on the one hand, *common* defamation, which in the balancing judgment may be considered to have succumbed to the protection of freedom of thought expression; on the other hand, forms of defamation such as to degenerate into hate crimes or otherwise into forms of incitement to violence, which cannot be tolerated in the EU legislative framework.

3.2.- *The French model*

Having exhausted the examination of the European conventional framework, we can begin to examine how individual states that have introduced *ad hoc* legislation on countering the spread of *fake news* fit within it.

The first model comes to us from the French legal system, within which, despite the provision of a robust arsenal of rules deemed appropriate to limit the dissemination of fake news, in balance with the right to freely express one's thoughts and freedom of

---

[62] "*Although the Contracting States are permitted, or even obliged, by their positive obligations under Article 8 of the Convention [...] to regulate the exercise of freedom of expression so as to ensure adequate protection by law of individuals' reputations, they must not do so in a manner that unduly deters the media from fulfilling their role of alerting the public to apparent or suspected misuse of public power.*" EDU Court, *Cumpănă and Mazăre v. Romania*, December 17, 2004, § 113.

[63] EDU Court, *Fatullayev v. Azerbaijian* judgment, April 22, 2010.

[64] EDU Court, *Katrami v. Greece* judgment, April 16, 2009.

the press, among which the *Loi du 29 juillet 1881 sur la liberté de la presse*[65] plays a central role, in 2018, as the elections for the renewal of the European Parliament approached, an intense debate developed on the advisability of introducing legislation specifically dedicated to the phenomenon of information manipulation.

The result was the presentation of *twin bills*[66] aimed at limiting the phenomenon of the spread of *fake news* in the electoral sphere; and this was because "*l'actualité électorale récente a démontré l'existence de campagnes massives de diffusion de fausses informations destinées à modifier le cours normal du processus électoral par l'intermédiaire des services de communication en* ligne."[67] .

Thanks in part to the use of the *procédure accélérée* requested by the government in March-both bills were passed in December 2018-the French legal system has now been enriched by *Loi organique n° 2018-1201 du 22 décembre 2018* and *Loi n° 2018-1202 du 22 décembre 2018*, both *relating à la lutte contre la manipulation de l'information*[68] .

This is, formally[69] , the only *criminal* law on *disinformation* provided by a European Union country, stipulating penalties of up to one year imprisonment and 75,000 euros in fines for violating transparency obligations imposed on IT platform operators.

---

[65] That in Art. 27 - as amended by *Ordonnance* n°2000-916 of September 19, 2000 - punishes the publication, dissemination or reproduction by any means of false news, invented, falsified or wrongly attributed to third parties in bad faith, if they disturb or are likely to disturb the public peace, with a fine of forty-five thousand euros, a penalty which, by reason of what is regulated in the second paragraph, shall be a fine of one hundred thirty-five thousand euros if the publication, dissemination or reproduction carried out in bad faith is of such a nature as to undermine the discipline or morale of the Army or to hinder the war effort of the Nation.

[66] These were, in particular, a *Proposition de loi organique* and a *Proposition de loi relative à la lutte contre la manipulation de l'information-it* should be noted, however, how, originally, the rubric mentioned *lutte contre les fausses informations*.

[67] *Proposition de loi relative à la lutte contre les fausses informations, n° 799, déposé(e) le 21 mars 2018.*

[68] The full text of which is available at: www.legifrance.gouv.fr.

[69] It should be mentioned in passing how the sanctioning severity of the German law raises more than one doubt as to whether it is actually *administrative* in nature, falling instead under the notion of *matière pénale* as defined by the EDU Court since the Grand Chamber judgment, June 8, 1976, *Engel v. Netherlands* (a definition later refined by *Öztürk v. Germany*, February 21, 1984, *Jussila v. Finland*, November 23, 2006; A. *and B. v. Norway*, November 15, 2016) On the point, F. MAZZACUVA, *La materia penale e il "doppio binario" della Corte europea, le garanzie al oltre delle apparenze*, in *Riv. it. dir. proc. pen.*, 4-2013, 1899 ff.; F. PALAZZO, *Il limite della* political question *fra Corte costituzionale e Corti europee. What is "substantially criminal,"* in M. Donini-L. Foffani (ed.), *La "materia penale" tra diritto nazionale ed europeo*, Turin, 2018, 7 ff.

The most interesting aspect of this legislation-beyond the choice to limit the law's operation to the election period alone[70] , which is considered a particularly *sensitive* time in the country's democratic life, so much so as to justify some restrictions on the right to freely express one's thoughts-is in the obligations[71] introduced against the operators of IT platforms[72] .

Prominent among these is that of providing the user with "*loyale, claire et transparente information*" on the identity of the natural person or the business name, registered office and corporate purpose of legal entities that pay a fee to the platform in exchange for promoting content related to a topic of general interest (*débat d'intérêt général*) [73] .

French legislation, too, has not gone without criticism, made in particular by the main trade union organization of French journalists (*Syndacat National des Journalistes - SNJ*), with reference to a somewhat *procedural* provision, namely that which allows the (civil) judge, at the instance of the public prosecutor, each candidate, party, political movement or anyone with an interest in the case, to take, in the three months preceding a

[70] The election period is identified as "*Pendant les trois mois précédant le premier jour du mois d'élections générales et jusqu'à la date du tour de scrutin où celles-ci sont acquises.*"

[71] Article 1 of Law 1202/2018 introduced some changes to Article 163-1 of the *Code électoral*, imposing several transparency obligations on the operators of information platforms. In particular, the platform must then provide the same user (2° de l'article L. 163-1) with *fair*, *clear*, and *transparent* information on the use of his or her personal data with regard to the promotion of an information content corresponding to a topic of general interest. Finally (3° de l'article L. 163-1) digital platforms have a duty to make public the total amount of remuneration received in exchange for the promotion of said content, if it exceeds a threshold that is set by decree. This information must be kept on a register made available to the public for electronic consultation, and failure to comply with these provisions is punishable by a one-year prison term and a 75,000 euro fine. Collective entities are also sanctioned: *Les personnes morales déclarées responsables pénalement, dans les conditions prévues à l'article 121-2 du code pénal, de l'infraction définie au premier alinéa du présent article encourent, outre l'amende suivant les modalités prévues à l'article 131-38 du même code, les peines prévues aux 2° et 9° de l'article 131-39 dudit code. L'interdiction prévue au 2° du même article 131-39 est prononcée pour une durée de cinq ans au plus et porte sur l'activité professionnelle dans l'exercice ou à l'occasion de laquelle l'infraction a été commise*.

[72] The operators addressed by the law are expressly identified by reference to Article L. 111-7 of the *Code de la consommation*, where it states: "*I.- Est qualifiée d'opérateur de plateforme en ligne toute personne physique ou morale proposant, à titre professionnel, de manière rémunérée ou non, un service de communication au public en ligne reposant sur : 1° Le classement ou le référencement, au moyen d 'algorithmes informatiques, de contenus, de biens ou de services proposés ou mis en ligne par des tiers ; 2° Ou la mise en relation de plusieurs parties en vue de la vente d'un bien, de la fourniture d'un service ou de l'échange ou du partage d'un contenu, d'un bien ou d'un service.*" Thus, the provision will apply to all search engines (e.g., *Google*), *online* review sites (e.g., *Tripadvisor*), exchange sites for goods or services (e.g., *Amazon* or *Airbnb*), and *social networks* (e.g., *Facebook* or *Twitter*).

[73] Paragraph 1 of Article 163-1 of the *Code électoral*, as amended by *Loi n° 2018-1202.*

general election, all proportionate and necessary measures to stop the deliberate, artificial, automatic and massive dissemination of false or tendentious information[74] .

The fact that in such cases, according to c. II of Art. 163-2 of the *Code électoral*, the judge must decide on a request for the removal of informational content within forty-eight hours of the notification of the complaint, has led French journalists to consider this law to be potentially the bearer of a liberticidal vision[75] , since in an extremely short time a decision will have to be made on investigations of long duration and great complexity, with the risk of weakening - even further - the position of professional journalists and exposing the secrecy of their sources to risk.

However, the issue, which was submitted to the *Conseil constitutionnel for* consideration, was rejected in *Decision* No. 2018-773 DC on December 20, 2018[76] .

3.3.- *The German model: the 'administrative' liability of* social network *operators in the NetzDG*

[74] This is the text of Article 163-2, c. I, of the *Code électorale*: "*Pendant les trois mois précédant le premier jour du mois d'élections générales et jusqu'à la date du tour de scrutin où celles-ci sont acquises, lorsque des allégations ou imputations inexactes ou trompeuses d'un fait de nature à altérer la sincérité du scrutin à venir sont diffused de manière délibérée, artificielle ou automatisée et massive par le biais d'un service de communication au public en ligne, le juge des référés peut, à la demande du ministère public, de tout candidat, de tout parti ou groupement politique ou de toute personne ayant intérêt à agir, et sans préjudice de la réparation du dommage subi, prescrire aux personnes physiques ou morales mentionnées au 2 du I de l'article 6 de la loi n° 2004-575 du 21 juin 2004 pour la confiance dans l'économie numérique ou, à défaut, à toute personne mentionnée au 1 du même I toutes mesures proportionnées et nécessaires pour faire cesser cette diffusion*".

[75] The full text of the SNJ press release titled *"Fake news: un projet de loi liberticideè,"* dated March 18, 2018, can be found at: www.snj.fr.

[76] *Decision* No. 2018-773 DC of Dec. 20, 2018, § 11, available at: www.conseil-constitutionnel.fr. The Conseil had been referred to the matter by a group of Senators and Deputies, who complained how this provision would result in an unproportionate compression of the freedom of manifestation of thought in political debate during the election campaign, due to the difficulty for the court to make a decision within forty-eight hours and to judge on the risk of altering the validity of a vote that has not yet taken place. According to the college, the provision of an emergency procedure to ensure the rapid cessation of conducts of dissemination of false information, such as to alter the genuineness of the vote, is in conformity with the Constitution, as it is limited in time, referred exclusively to online communication services-that is, to those media through which it is easier to carry out massive and coordinated manipulation actions-and limited by law to only inaccurate or misleading news, such as to affect the imminent exercise of the right to vote. In the *Conseil*'s view, opinions, parodies, partial inaccuracies or simple exaggerations are not open to scrutiny by the court, but only those news items whose objective falsity can be demonstrated. That being said, in view of the fundamental role played by the right to freely express one's thoughts in the period leading up to a general election, the legislature, in allowing the judge to prescribe all proportionate and necessary measures to stop the dissemination of illicit content, has also required him to take those measures that are the least harmful to freedom of expression and communication.

The second model to consider is the one adopted in the Federal Republic of Germany, where the *Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken*[77] , better known as *Netz DG* or *Facebook Act, was* introduced.

The events leading up to the enactment of this law are of equal interest to its content, describing a clear failure of the network's self-regulatory mechanisms.

In the face of the commitment made by the main *media companies* to introduce internal reporting mechanisms to ensure the removal of offensive content within twenty-four hours, an independent analysis of the removal of offensive content by the operators of the main *social networks*, conducted on behalf of the federal government by an organization for the monitoring of child protection laws (*Jugendshutz.net*), yielded largely negative results[78] ; so that, partly as a result of the further awareness gained regarding the offensive potential of the massive dissemination of *fake news* through the network, following the course of the U.S. presidential election campaign[79] , the *Netz DG* was presented in March 2017 and approved five months later, despite the firm opposition of *social network* operators and some sectors of German civil society, concerned about the possible compression to the freedom of manifestation of thought[80] .

---

[77] The full text of the law is available at: www.bmjv.de. For an Italian translation, see *MediaLaws - Rivista dir. media*, 1-2017, 185 ff. In the remainder of the discussion, we will refer to this text.

[78] During two observation periods of eight weeks each, conducted from July to August 2016 and January to February 2017 respectively, *Jugendschutz.net* investigated content that violated Article 130 (incitement to hatred and Holocaust denial) and Article *86a* (use of symbols of unconstitutional organizations) of the *Strafgesetzbuch*, and the Youth Protection Act. Based on 200 reported contents per tested platform, *Facebook* removed 39 percent, *YouTube* 90 percent, and *Twitter* 1 percent. Looking only at content removed within twenty-four hours of being reported, the percentages dropped to 31% for *Facebook*, 82% for *YouTube*, and 0% for *Twitter*. v. *Löschung Rechtswidriger Hassbeiträge Bei Facebook, YouTube Und Twitter: Ergebnisse Des Monitorings von Beschwerdemechanismen Jugendaffiner Dienste" (Jugendschutz.net, 2017)*, at www.astrid-online.it. Hence, according to the *Netz DG* proponents, it emerged how *social network* operators, despite their stated commitments and at the expense of the fact that technology allowed them to intervene, had failed to live up to their task of self-regulation (Literally: "*Die Anbieter der sozialen Netzwerke haben eine Verantwortung, der sie gerecht werden müssen. Angesichts der Tatsache, dass das bisherige Instrumentarium und die zugesagten Selbstverpflichtungen der sozialen Netzwerke nicht ausreichend wirken und es erhebliche Probleme bei der Durchsetzung des geltenden Rechts gibt, bedarf es der Einführung von bußgeldbewehrten Compliance-Regeln für soziale Netzwerke, um effektiv und unverzüglich gegen Hasskriminalität und andere strafbare Inhalte im Netz vorgehen zu können*").

[79] This is stated in the preamble to the bill that introduced *Netz DG*: "*ach den Erfahrungen im US-Wahlkampf hat überdies auch in der Bun- desrepublik Deutschland die Bekämpfung von strafbaren Falschnachrichten ("Fake News") in sozialen Netzwerken hohe Priorität gewonnen.*"

[80] See for all: G. NOLTE, *Hate-Speech, Fake-News, das "Netzwerkdurchsetzungsgesetz" und Vielfaltsicherung durch Suchmaschinen*, *Zeitschrift für urheber- und medienrecht*, 7-2017, 552, 554.

From a subjective standpoint, the law applies (§1.1) *to telecommunications service providers that for profit operate Internet platforms designed to enable users to share any kind of content or make it accessible to the public (social networks)*, with the express exclusion from its scope of application of those platforms *that offer journalistic or editorial content,* which are not considered, as well as those platforms that are intended for individual communication or dissemination of specific content.

In addition to *qualitative requirements*, the law includes a *quantitative* requirement, applying (§ 1.2) only to *social networks* with more than two million registered users in Germany.

The obligations that the law places on *social network* operators are of two types: on the one hand, a duty to report on complaints received about the publication of illicit content is introduced (§. 2); on the other hand, a duty to handle said complaints (§3) is provided, such as to ensure that access to manifestly infringing content is removed or blocked within twenty-four hours of receipt of the report, unless the *social network* has agreed with the competent judicial authority on a longer period for the removal or blocking of manifestly infringing content (§3.2.2), or the removal or blocking without delay of all infringing content, within seven days of receipt of the report (§3.2.3)[81] .

As anticipated, the penalties for violation of these duties are extremely severe: for the operator who, in violation of § 3, sub-section (4), second sentence, fails to correct an organizational deficiency or eliminate it in a timely manner, a fine of up to five hundred thousand euros may be imposed; in the case of violation of the duty to prepare the report referred to in § 2, sub-section (1), or to prepare it incorrectly, incompletely, or late, as well as in the case of violation of the duty to remove unlawful content referred to in § 3 sub-sections (1) and (4), a fine of up to five million euros shall be imposed.

Finally, there is a form of extraterritoriality of German jurisdiction, the sanctions provided for in *Netz DG* being applicable even to acts not committed in the *FRG*.

Again, the criticism was not limited to assessments of political expediency, but concerned aspects of constitutionality on the domestic and European levels, both with

[81] According to the law, the seven-day time limit may be exceeded if: a) the decision on the unlawfulness of the content depends on the falsity of a statement or clearly depends on the factual circumstances; in such cases, the social network operator may give the user an opportunity to respond to the report before a decision is made; b) the social network refers the decision on unlawfulness to a self-regulatory body recognized under subsections (6) and (8) within seven days of receiving the report, agreeing to comply with the decision of such body.

reference to the *topos of* freedom of thought and with reference to the limits that the *E-commerce* Directive imposes on member states.

On the domestic side, it was noted how the first paragraph of Article 5 of the *Grundgesetz* enshrines for everyone "*the right to freely express and disseminate his or her opinions in words, writings and pictures, and to inform themselves without hindrance from sources accessible to all*," as well as guaranteeing "freedom of the *press and freedom of information through radio and film*" and providing for the prohibition of censorship.

As is typical of the continental constitutionalist tradition, it is the Fundamental Law itself that provides for certain limits to the exercise of these freedoms, which the second paragraph of Article 5 identifies "*in the provisions of general laws, legislative norms concerning the protection of youth and the right to respect for the honor of the person.*"

The problem in this case, as in others we are about to examine, is to establish with certainty which opinion can be freely expressed and which, on the other hand, can be restricted because of its conflict with mandatory rules of law.

In the case of the *NetzDG*, censorship for someone else's thought is not directly established, but a penalty is imposed for the omissive conduct-or even unjustified delay-of the operator of a social network, against whom a duty to remove illegal content is established.

It has thus been observed[82] how this system induces intermediaries to adopt a prudential evaluation criterion, removing any dubious content; an evaluation criterion also favored by the absence of the provision of penalties for those who unduly remove lawful content, in accordance with the spirit of a law whose purpose is not to prevent unlawful decisions, but rather to prevent the dissemination of unlawful content[83] .

Some doubts have also been raised with reference to the requirement of *proportionality, to which* Section 5(2) of the GG refers, as interpreted by the *Bundesverfassungsgericht to* mean that exceptions to the principle of full freedom of

[82] M. LIESCHING, *Die Durchsetzung von Verfassungs- und Europarecht gegen das NetzDG - Überblick über die*
wesentlichen Kritikpunkte, in *MultiMedia und Recht*, 1-2018, 27.
[83] K.-H. LADEUR-T. GOSTOMZYK, *Gutachten zur Verfassungsmäßigkeit des Entwurfs eines Gesetzes zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken*, in www.bitkom.org, 2017,81.

thought are permissible only if they pursue a legitimate aim and are appropriate, necessary and, strictly speaking, proportionate to that aim[84] .

Finally, the transnational location of the various actors involved in the *NetzDG*'s enforcement mechanism also makes it necessary to assess the consistency of its provisions with the E-Commerce Directive[85] , where in Art. 3, § 2, it is stipulated that *Member States may not, for reasons falling within the regulated field, restrict the free movement of information society services from another Member State*, without prejudice (Art. 3, § 4, lett. a(*i)* the possibility for member states to introduce exemptions for reasons of public policy, in particular for the work of prevention, investigation, detection and prosecution in criminal matters: such as the protection of minors and the fight against incitement to racial, sexual, religious or ethnic hatred, or (Art. 3, § 4(a)(*ii*) relating to a particular information society service that is detrimental to the objectives referred to in (i) or poses a serious and grave risk of harm to those objectives.

3.4. *Anti-fake news laws and cyberspace law in authoritarian democracies and Asian People's Republics.*

If we were to draw an ideal map of the jurisdictions that have adopted criminal laws to crack down on the spread of *fake news*, the Asian quadrant would be vastly in the majority.

For obvious geo-political reasons, it is useful to begin with an examination of China's criminal justice system, which is considered one of the strictest in the world, not only because of the wide and indiscriminate use of the death penalty.

As of 2015, after a number of egregious cases of dissemination of fake news had already occurred in the early 2000s[86] - qualifying as true forerunners of modern *fake news*

[84] See on this point V. CLAUSSEN*, Fighting hate speech and fake news. The Network Enforcement Act (NetzDG) in Germany in the context of European legislation*, in *MediaLaws - Journal dir. media*, 3-2018, 122.

[85] Directive 2000/31/EC of the European Parliament and of the Council of June 8, 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market ("E-Commerce Directive"), the full text of which, in Italian, is available at: www.eur-lex.europa.eu.

[86] For example, in 2008, the spread of false rumors about the finding of worms in tangerines grown in the Sichuan region severely harmed the area's farmers; just as in 2011, after the Fukushima nuclear accident, rumors were spread on the QQ *social network* about the power of salt to protect against radiation, which

- an incriminating case was introduced in the second paragraph of Art. 291 of the Criminal Code, an incriminating case punishing with imprisonment up to three years anyone who produces false information referring to dangerous situations, epidemics or catastrophes, or intentionally spreads such news, knowing that it is fake, in such a way as to cause a serious danger to public order; if such conduct causes serious consequences, the penalty is imprisonment from three to seven years.

The norm - which according to Western classification criteria must be considered as a crime of presumed danger, placed to protect public order *sub forma of* public tranquility, to be guaranteed against those who intend to arouse panic in the population by spreading false news about events capable of exposing the community to danger - There seems to be characterized by an extremely broad legal type - difficult to attribute connotation of precision to the notion of *dangerous situation*, to which it refers - but not such that it could encompass *fake news* with defamatory content or even without injurious effects, which would in any case be incapable of disturbing public order. Reason why in China, according to the Report "*Initiatives to Counter Fake News*" compiled by the *Law Library of Congress*, despite strict regulation of the *media* and the Internet, *fake* news-what local *media* refer to as *online* rumors-is nonetheless widespread[87] ; as evidenced, among other things, by the Chinese government's decision to disseminate, as of August 2018[88] , through its own platform - called *Piyao*, a word that in Mandarin means "*refutation of* rumors"[89] - *real* news as it comes from state-run media or otherwise local newspapers controlled by the Communist Party or government agencies.

Finally, consider that according to the *cybersecurity* law adopted by the People's Republic of China in 2016, not only are numerous *online* activities prohibited, including the production or dissemination of fake news such as to disturb the economic and social order, but provision is also made for operators of sites that provide news publishing and dissemination or instant messaging services, the obligation to register their users with real names, as well as to provide such data to the government if requested to do so, with the

triggered a veritable *hunt for the* mineral by large masses of Chinese citizens. see M. REPNIKOVA, *China's Lessons for Fighting Fake News*, in www.foreignpolicy.com, September 6, 2018.

[87] According to a report released by the *Wechat* platform, more than eighty-four thousand *fake news stories were* disseminated within it in 2019, read by two hundred and ninety-four million users. v. *2018 Report on Managing Online Rumors Published, 774 Institutions Refuted Rumors on Wechat*, in www.people.cn, January 18, 2019.

[88] V. *China launches platform to stamp out 'online rumors'*, in www.reuters.com, August 29, 2018.

[89] *www.piyao.org.cn.*

provision, in case of failure to comply, of administrative penalties of up to seventy-five thousand dollars[90] .

More recently, the Chinese government announced that it would introduce new oversight requirements on *deepfake* technology providers. Specifically, they would like to regulate *deep synthesis internet information services*, a notion under which any technology that generates text, images, audio, video or virtual scenes based on *deep learning*[91] .

According to this draft regulation, *deepfake* service providers would be required to verify the identity of every user and actively monitor the results of their services for potential violations. Providers would also have to "*place marks that do not interfere with the user experience*" to make all *deepfake* content identifiable and traceable[92] .

Thus succinctly reconstructed, the Chinese experience reveals several aspects of great interest.

While not foregoing criminal leverage-typical of non-democratic systems-the Chinese government is fully aware of the role of *gatekeepers* in network regulation and provides forms of *law enforcement* aimed at targeting service providers.

At the same time, forms of repression of manipulative conduct-whether based on the spread of *fake news* or *deepfakes-are* part of a broader project of regulating *cyberspace*, which constitutes-as has long been noted-the new battleground between the superpowers of the 21st Century[93] .

To see how the Chinese model is leading the way, one need only travel to Malaysia, where an *Anti-fake news act*[94] had been passed in 2018, which stipulated, for those who publish or share *fake news*, a penalty of up to six years' imprisonment and a fine of five hundred thousand *ringgit*[95] .

The law - extremely controversial - was finally repealed in October 2019[96] , following the defeat in the general elections[97] of the majority that had supported its introduction, and after the combination of the severity of the sanctioning response and the absolute lack of taxability and precision of the incriminating case - such as to compress

---

[90] *PRC Cybersecurity Law* (passed November 7, 2016 and effective June 1, 2017), in*:* www.npc.gov.cn.
[91] https://www.openaccessgovernment.org/chinas-deepfake/128860/
[92] https://www.protocol.com/bulletins/china-deepfake-regulation
[93] KAI-FU LEE, *AI Superpowers. China, Silicon Valley and the new world order*, Boston, 2018.
[94] The full text can be found at: www.federalgazette.agc.gov.
[95] That is roughly equivalent to one hundred and fifteen thousand euros.
[96] v. *Malaysia parliament scraps law criminalizing fake news*, in www.aljazeera.com, October 10, 2019.
[97] See on this point the article by M. MONTI, *Chronicle of the enactment and repeal of the Malaysian Anti-Fake News Act*, in *MediaLaws - Journal dir. media*, 3-2018, 435 ff.

any form of dissent against the government and other state authorities[98] - had led to strong opposition from nongovernmental organizations engaged in the defense of human rights, which noted that through this law a kind of state censorship would be realized[99] .

The censorious intentions of the rule were quite evident from the definition of *fake news* adopted in it, which included "*news, information, data and reports, which are or are wholly or partially false, in the form of content, images or sound recordings or in any other form capable of insinuating words or* ideas."[100] ; so much so that the definition of the *Malaysian gag*, elaborated in doctrine, seemed to hit the nail on the head[101] .

After all, Malaysia has long been considered among the countries in which freedom of thought and press freedom are least guaranteed, so much so that the *World Press Freedom Index* placed it at position 123/180[102] in 2019, while the compilers of *the Freedom House Freedom of the Press report* gave it the rating of *not free* in 2017[103] and this is not only because of the recent *Anti-fake news* law, but rather because of the already strict regulatory arsenal that, starting with the *Sedition Act* of 1948, through the *Official Secrets Act* of 1972, and finally the *Communications and Multimedia Act* of 1998, outlines an extremely problematic environment for both the exercise of the journalistic profession and for that pluralism in the circulation of ideas that is the basis of any truly democratic system.

Equally radical was the approach taken by the Socialist Republic of Vietnam, which in June 2018 passed a law[104] -which went into effect on January 1, 2019-regarding *cybersecurity*, which introduces a number of limitations and obligations on *internet*

---

[98] The extent of the rule's enforcement reach is demonstrated by the first documented case of *Anti-fake news Act 2018* enforcement, which involved a Danish citizen, Salah Salem Saleh Sulaiman, who was indicted for spreading fake news by posting a video on the *YouTube* platform, in which he accused Malaysian police of taking more than fifty minutes to intervene after a firefight in which a Palestinian teacher was killed was reported. The police claimed they intervened after only eight minutes, and Sulaiman was sentenced to a ten thousand ringgit fine (about two thousand five hundred euros), converted to one month's imprisonment due to his inability to provide payment. v. *Danish national first to be convicted under Malaysia's fake news law*, in www.reuters.com, July 28, 2018.

[99] *Malaysia accused of muzzling critics with jail term for fake news*, in www.theguardian.com, March 26, 2018.

[100] The translation is ours. This is the original text of the rule, contained in Article 2, c. 1 of the *Anti-fake news act 2018*: "*any news, information, data and reports which are wholly or partly false, whether in the form of features, visuals or audio recordings or in any other form capable of suggesting words or ideas.*"

[101] A. MAZZIOTTI DI CELSO, *From the First Amendment to the Malaysian Gag. Fake news, freedom of expression and the reversal of traditional political categories*, in *MediaLaws - Journal dir. media*, 3-2018, 90 ff.

[102] However, in 2018, the year the law we are discussing was passed, Malaysia was ranked 145/180.

[103] *Freedom of the Press 2017: Malaysia, Freedom House*, at www.freedomhouse.org.

[104] The full text of which is available at: *www.economica.vn.*

service providers, who are required to disclose user data to the government, thus allowing identification of who has published certain content.

Although this is not specifically a law on *fake news*, there are numerous normative references to the phenomenon we are examining.

First, in Art. 8 of the law, where the conducts for which the use of *cyberspace* is prohibited are listed: the *distortion of history (distorting history*); *denying revolutionary achievements (denying revolutionary achievements*); *destroying the national solidarity* block (*destroying the national solidarity block*); insulting religion, discriminating against gender, and engaging in racist acts; spreading false information causing unrest among the citizenry, damage to socio-economic activities, hindering the operation of state agencies or persons managing public property, or harming the rights and legitimate interests of other entities, organizations or individuals[105] .

Even more analytical is the listing of conduct prohibited by Article 16, headed "*Prevention of and dealing with information in cyberspace with contents being propaganda against the Socialist Republic of Vietnam; information contents which incite riots, disrupt security or cause public disorder; which cause embarrassment or are slanderous; or which violate economic management order.*"

In the latter case, we are dealing with a crasis between protection paradigms typical of the authoritarian regimes of the twentieth century[106] and the protection of market freedom[107] , in deference to the *Socialist market economy* (SME) paradigm[108] adopted in China since the early 1990s and now extended to many of the countries within its sphere of influence.

---

[105] The translation is Nostra, this is the original text: "*Providing false information, causing confusion among the Citizens, causing harm to socio- economic activities, causing difficulties for the operation of State agencies or of people performing public duties, or infringing the lawful rights and interests of other agencies, organizations and individuals*."

[106] The conduct of *Distortion or defamation of the people's administrative authorities*, i.e., *Insulting the [Vietnamese] people, the national flag, national emblem, national anthem, great men, leaders, famous people or national heroes*, is sanctioned, as is the prohibition of using *cyberspace Calling for, mobilizing, inciting, threatening, or embroiling a mass/crowd of people to disrupt or oppose people [officials] conducting their official duties, or obstructing the activities of agencies or organizations causing instability to security and order.*

[107] The prohibited conduct also includes "*Invented or untruthful information about products, goods, money, bonds, bills, cheques and other valuable papers*"; "*Invented or untruthful information in the sectors of finance, banking, e-commerce, e-payment, currency trading, capital mobilization, multi-level trading and securities.*" These are, as evident, conducts that do not offend the interests of the state - as would have been logical to expect in the context of a state that still refers - formally - to the principles of real socialism - but rather economic interests of private individuals, which are deemed to take precedence over the freedom to freely manifest one's thoughts.

[108] v. X. DING, *The Socialist Market Economy: China and the World* Science & Society, Vol. 73, No. 2, *China: Socialism, Capitalism, Market: Why Not? Where Next?*, Apr., 2009, 235 ff.

However, the concern to extend the laws in force in the world of atoms to the world of *bits* is not the prerogative of countries influenced by the Chinese economy.

In May 2019, Singapore, a former British colony in which staunch adherence to the capitalist system is accompanied by a political regime attributable to authoritarian democracy[109] , also passed a law criminalizing the dissemination of false information *online*.

The law, which went into effect on Oct. 2 of that year, introduced one of the most stringent information legislation in the world, both in terms of the breadth of punishable conduct and in terms of penalty treatment.

According to Art. 7 of the *Protection from online falsehoods and manipulation act 2019*, any person in any part of the world who disseminates objectively false news (*false statements of fact*) concerning Singapore, such as to prejudice the security of the country or a region thereof, public health or safety, public tranquility, or public finances, or such as to impair friendly relations between Singapore and other countries of the world, shall be punished or influence the results of a presidential or parliamentary election, incite feelings of enmity or hatred among groups of people, or such as to diminish public confidence in the activity of any organ or power of the State (*any duty or function of, or in the exercise of any power by, the Government, an Organ of State, a statutory board, or a part of the Government, an Organ of State or a statutory board*).

The punishments, listed in Article 7, § 2, are both pecuniary in nature and deprivation of liberty and can reach, in the case of acts committed by natural persons, up to ten years' imprisonment and, for legal persons, up to a million dollar fine.

The severity of Singapore's legislation is also evident by examining the provisions (Articles 10, 11, 12) of *law enforcement*, where the government is given the power to address to *social network* operators a correction *direction* (request for correction) or a stop communication *direction (*request to stop communication) respectively aimed at correcting an allegedly false news or removing it without delay "*even if the recipient has no reason to believe that such news is false.*"

Therefore, it will be the government - but the power to act is also given to an individual minister - who will decide which news - ontologically *true* - can be known by

[109] For an in-depth study, see the study by P. KHANNA, *The Rebirth of the City-States*, Rome, 2017, 36 ff.

Singaporean citizens (voters) and which news - to be considered without possibility of appeal as *fake news* - should not be spread[110] .

3.5.- *Russian legislation, between* disinformatja*, troll factories and criminal repression of dissent.*

No research dealing with *fake news* can be said to be complete without considering the Russian experience.

This is not just a matter of *tradition, although* it is still true that propaganda has been an essential element of Soviet power since the October Revolution, although particularly used in Stalin's time.

In the Soviet Union, regime propaganda frequently relied on conspiracy theories: as Hannah Arendt observed, "*during the Stalin era there was a real accumulation of alleged conspiracies; the discovery of a new conspiracy did not cause the* previous *one to be set aside.*"[111] .

Brief.

Manipulation of the masses has always been a characteristic feature of successive regimes in Russia, as even the most recent news from the Ukrainian front shows.

After all, Vladimir Putin's entire governing experience is intertwined with the issue of *fake news*, of which Russia turns out to be a major exporter.

In particular, following the publication of the *Mueller Report there* is evidence of how the Russian government interfered in various ways with the 2016 U.S. presidential election, not only through the massive preparation of *misinformation* campaigns through *fake* groups and profiles, but also through *hacking* attacks as a result of which confidential documents of the *Democratic National Committee* were stolen, then disseminated through various *web* platforms, including *WikiLeaks*[112] .

Not only that.

---

[110] The law is having its first effects: on November 30, 2019, *Facebook* announced that it had added, in deference to the provisions of the *Protection from online falsehoods and manipulation act 2019,* in its Singapore version, a footnote to a November 23 post by Alex Tan, author of the *State Times Review*, a *blog* often critical of the government. The text of the indication reads, "*Facebook is legally required to tell you that the Singapore government says this post has false information*." v. M. BUTCHER, *Facebook bowed to a Singapore government order to brand a news post as false*, in www.techcrunch.com, Nov. 30, 2019.

[111] H. ARENDT, *The Origins of Totalitarianism,* cit., 485, nt. 29.

[112] *The Mueller Report. The final Report of the Special Counsel into Donald Trump, Russia and collusion*, New York, 2019, 38.

Russia is home to the *Internet Research Agency* (IRA), a St. Petersburg-based company known to be one of the world's leading *Troll Farms*[113] , which has been conducting large-scale *social* media campaigns since mid-2014 using fake profiles, ostensibly owned by U.S. citizens, to spread propaganda messages and *fake news*[114] .

To accomplish this, the IRA had hired dozens of employees - called *specialists* - assigned to manage fake profiles and pages created *ad hoc* to appear to be referable to U.S. entities or institutions[115] .

More recently, the activities of Russian agencies in the context of the *Brexit* referendum campaign, and more generally aimed at influencing the British political system, have been the subject of an investigation conducted by the Intelligence *and Security Committee* (ISC)-the parliamentary oversight body for intelligence services in the United Kingdom-whose findings have not yet been released[116] .

Pending these further investigative results, however, it does not seem doubtful that within the territory of the Russian Federation, entities dedicated to the production and dissemination of massive disinformation campaigns have operated and are still operating over the past few years, at which numerous *content creators* act, whose job is to create and manage as many fake *accounts as* possible, with the aim of influencing the debate and affecting consensus-building in other countries of the world, during electoral competitions[117] .

---

[113] Cf. I. LAPOWSKY, *Facebook May Have More Russian Troll Farms to Worry About*, in www.wired.com, September 8, 2017.

[114] *The Mueller Report,* cit., 53.

[115] None of the major *social networks* remained immune to this disinformation activity, although the Russian *Trolls*' activity was mainly concentrated on *Facebook*, where - in addition to the combined action of fake profiles and *owl* pages - over a hundred thousand dollars were invested in advertisements. *Ibid*, 56-57.

[116] *Ministers questioned on the publication of the Intelligence and Security Committee's report into Russia*, November 5, 2019, in: https://www.parliament.uk/business/news/2019/november/ministers-questioned-on-the-publication-of-the-intelligence-and-security-committees-report-into-russia/

[117] According to the declassified report by the *United States Office of the Director of National Intelligence,* titled *"Background to "Assessing Russian Activities and Intentions in Recent U.S. Elections,"* dated Jan. 6, 2017, "*Russia's propaganda machine, consisting of its domestic media apparatus, tools aimed at a global audience-such as RT and Sputnik-and a network of para-governmental trolls, contributed to the disinformation campaign by serving as a platform for the Kremlin's messaging to Russia and international audiences.*" (Ibid., 3). The full text of the report is available at: https://www.dni.gov/files/documents/ICA_2017_01.pdf.

Nonetheless, in March 2019, Russia adopted two *anti-fake news* laws, amending the previous legislation on accuracy of information and providing administrative fines for anyone who spreads such news[118] .

Formally, in Russia, the Information Law[119] provides for freedom of manifestation of thought, unless it is aimed at propaganda of war, incitement to national, racial or religious hatred or enmity, as well as any other information from the dissemination of which criminal or administrative liability arises.

However, the government's system of control over the dissemination of information has always been very penetrating, as evidenced by the fact that if news is distributed through a news aggregator, the owner of the aggregator-and only Russian individuals or legal entities can be-is considered responsible for verifying the veracity of the facts of social significance published by them. In any case, state agencies have the right to apply to the federal executive body that exercises control and supervisory functions in the field of mass media, mass communications, and information technology (*Roskomnadzor*), to take the necessary measures to stop the distribution of such information, while, if a news dissemination in violation of the law takes place, the Attorney General of the Russian Federation or his deputies may petition the *Roskomnadzor, with a* request to stop the distribution of such information.

With the introduction of the March 2019 *anti-fake news* legislation, media enforcement powers have become even more penetrating: fines have been enacted for knowingly disseminating fake news, defined as "*socially significant false information distributed in the form of truthful messages, such as to pose a threat to life, health or property, or such as to create possible mass violations of public order or safety, or which may hinder transportation, the operation of infrastructure, lending institutions, communication lines, industry and* energy *companies*."[120] .

Having revamped the previous law on information, the mechanism for reporting and removal of such news is similar to the one we have just given an account of: the level of reliability of the information will have to be assessed by the Attorney General's Office,

---

[118] These are, respectively, Federal Law No. 31-FZ of March 18, 2019, amending Article 15-3 of the Federal Law on Information, Information Technology and Information Protection (available at: http://publication.pravo.gov.ru/Document/View/0001201903180031) and Federal Law No. 27-FZ of March 18, 2019, amending the Code of Administrative Violations (available at: http://publication.pravo.gov.ru/Document/View/0001201903180021? index=1&rangeSize=1).

[119] Federal Law No. 149-FZ on Information, Information Technology and Information Protection, July 27, 2006 (available at: http://pravo.gov.ru/proxy/ips/? docbody&nd=102108264).

[120] Law amending Article 15-3 of the Information Law, § 1 (our translation).

which is responsible for forwarding to the *Roskomnadzor the* request to take measures to restrict access to it.

As for penalties, the *anti-fake news* legislation refers back to the Code of Administrative Violations, where it provides for fines that, in the most serious cases-that is, those in which perpetrators of the offenses are collective entities-can reach a maximum of 1.5 million rubles (just over 21,000 euros).

The approval of the discipline of combating the dissemination of fake news through the web took place despite numerous critical remarks by the institutions involved - including the *Roskomnadzor* - concerned about the economic and organizational aspects that the reform would entail, with respect to the possible compression of subjective rights not particularly rooted in the legal tradition of that country[121] .

It cannot escape notice, moreover, how the option pursued by the Russian legislature bears strong similarities with the German model[122] , with which it shares the use of administrative fines, differing, however, in that they are applied directly to the author of the dissemination of *fake news*, rather than providing for a liability of the operator of the *social network* platform through which the dissemination takes place.

Nevertheless, upon closer examination, there are very few similarities with the German discipline, which is characterized by a careful balancing of the interests at stake.

What makes the Russian model far more similar to those adopted in authoritarian democracies or authoritarian regimes *tout court in* the Asian quadrant, than to those adopted in systems that are extremely careful to strike a proper balance between the different fundamental rights that come into play when dealing with *misinformation*, is the context in which this legislation-apparently consistent with the principle of proportionality-comes to be cast.

---

[121] A concern that has instead been expressed by some NGOs, including *Free Speech Association*, *St. Petersburg PEN* and *PEN Moscow*, which have noted that through the *fake news* laws the government has introduced tools to control freedom of expression through the Internet. v. Russia: *New laws threaten freedom of expression and media freedom*, in www.peninternational.org, April 1, 2019.

[122] In Germany, the so-called *Network Enforcement Act* or *NetzDG* (*Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken)* has been in force since 2018, which imposes an obligation on major digital platforms-specifically, on platforms boasting more than two million users-to promptly remove, following a specific complaint, any content that is qualified as unlawful by the German Criminal Code. The offenses covered by the legislation include, in particular, incitement to hatred and certain defamatory offenses, as well as cases in which content is published that is likely to constitute a danger to constitutional order or national security. On this point, see, most recently: M. GALIMBERTI, *Disinformation: regulatory and policy profiles*, in *Dir. pen. cont. riv. trim.*, 4/2021, 272.

Russia, in fact, has long been considered *not free* by reports compiled by the NGO *Freedom House*, reporting a score of 5/40 when it comes to respect for political rights and 3/16 when it comes to the protection of freedom of expression[123] .

On the other hand, suspicious disappearances or murders of independent journalists have not been uncommon over the past few years, occurring even after the outcry raised in the international community by the killing of investigative journalist Anna Politkovskaya, who had long been engaged in highly critical investigations at odds with the official truths disseminated by the Putin government[124] .

The strongly repressive attitude with respect to the possibilities of expanding civil liberties through the use of digital media is demonstrated, most recently, by a recent and much-discussed law called by the Western *media* "*sovereign Internet law,*" by which the government, in case of emergency can disconnect the connection to the network, preventing access to Internet sites located outside of Russia, and this with the stated purpose of limiting the possible effects of a possible *cyber-attack*, although it is not difficult to believe that this is yet another instrument of censorship, entrusted to the absolute discretion of those who hold executive power[125] .

An examination of Russia's preventive-repressive order regarding *fake news* would not be complete without considering what happened in conjunction with the February 2022 invasion of Ukraine.

Immediately after the launch of what is referred to in Moscow and Beijing as a "*special military operation*," the Duma approved a new *fake news law* that provides for the punishment of imprisonment of up to fifteen years-in addition to severe fines-for those who *knowingly* spread false information about all Russian state actors operating abroad.

At the same time as this further tightening of the sanctions framework, the Russian government announced-but, at the time of writing, not yet implemented-the disconnection from the *global* network with a switch to the autarkic RuNet, in execution of the previously mentioned *sovereign Internet law*.

---

[123] In the chapter devoted to the Russian Federation in *Freedom Report 2019,* we read, "Although the Constitution provides for freedom of speech, some laws with very vague content regarding the suppression of extremist activities grant the authorities great discretion in suppressing any speech, organization or activity that lacks official support. The government controls, either directly or through state-owned and or trustworthy companies, all national television networks and many radio and print media outlets, as well as most of the advertising market" (our translation). *Freedom in the world 2019*.

[124] A. POLITKOVSKAYA, *Proibito parlare*, Milan, 2007; ID. *Putin's Russia,* Milan, 2015.

[125] The law, which some say introduced a *digital iron curtain*, was justified with the need to defend Russia from *cyber attacks*. see E. SCHULZE, *Russia just brought in a law to try to disconnect its internet from the rest of the world, November* 1, 2019, in www.cnbc.com, November 1, 2019.

4. Conclusions

The picture that emerges from studying the legislation of authoritarian democracies thus calls for careful reflection.

It would be simplistic-and epistemologically incorrect-to sophistry on the *post hoc ergo propter hoc* model, judging legislation negatively solely by reference to the political context in which it was enacted.

Nonetheless, we cannot help but observe how the introduction of criminal sanctions aimed at punishing the conduct of spreading false news through the network, including through the provision of draconian penalties for the operators of social networks or information sites, assigning to government bodies the task of determining the true and the false - *i.e. what* can be known and debated by citizens and what must be kept from them - is completely incompatible with the democratic paradigm, which sees the freedom to freely express one's thoughts as one of the fundamental pillars of its conceptual edifice.

Brief: punishing those who produce and disseminate *fake news* is not in itself illegitimate, but the extreme difficulty of striking a precise balance between the interests at stake must be borne in mind, not least because of the centrality of freedom of expression in the European legal system.

This leads us to reiterate what we have observed in the past.

The problem of digital consensus manipulation, while touching on *private* interests[126] is characterized by a prevailing political dimension.

For this reason, legislative action is needed, possibly at the EU level, where something seems to be on the way to change with the proposed regulation "*on a Single Market For Digital Services (Digital Services Act)*," with which the European Commission proposes to update-twenty years after the *e-Commerce* Directive[127] -in order to create a secure digital environment, predictable and reliable, such as to ensure respect for the fundamental rights guaranteed by the Charter of Fundamental Rights of the

[126] See on this point the extensive study by A. VISCONTI, *Reputation, dignity, honor,* Turin, 2018, as well as the considerations more recently expressed by S. DE FLAMMINEIS, *Criminal law, collective legal goods in the challenge of fake news: principle of offensiveness and emergencies*, in *Sist. pen.*, 6/2020, 139-140.
[127] Directive 2000/31/EC of the European Parliament and of the Council of June 8, 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market (Directive on electronic commerce).

European Union and, in particular, the right to manifest thought, to information, to conduct one's business freely and not to be discriminated against.

To do this, the DSA identifies a series of obligations-accompanied by a strict penalty apparatus-addressed to providers of intermediary services (providers *of intermediary services*), imposing on them obligations proportionate to their role, size and impact on the digital ecosystem[128] .

This is a first step, but one that should be followed by initiatives on the level of individual EU member states as well.

Initiatives that, in our opinion, should exclude the introduction of criminal regulations, reaffirming even in the Infosphere the fundamental value of the principle of *extrema ratio*[129] .

Precisely because the current political-criminal temperament pushes toward pan-penalization, whenever possible the temptation to implement a now plethoric punitive system should be avoided, preferring to prune a few dry or minor branches - according to an effective metaphor by Francesco Palazzo[130] - rather than studying new grafts in the laboratory.

In a historical phase in which the crisis of the liberal[131] - and constitutionally oriented[132] - paradigm of criminal law is now being openly discussed, and in which, as observed by Filippo Sgubbi, "*public ethics is being transformed into* criminal *law*"[133] ,

---

[128] For a more in-depth examination of DSA, see: L. D'AGOSTINO, Disinformation and platform liability. Triggering obligations and compliance measures, in *Dir. pen. cont. riv. trim.*, 4/2021, 291 ff.

[129] See, among others: F. MANTOVANI, *Diritto penale,* V Ed., Padova, 2007, XL et seq.; F. BRICOLA, *Carattere sussidiario del diritto penale e oggetto della tutela*, in *Studi in onore di Delitala*, Vol. I, Milano, 1984.
, 99; T. PADOVANI, *La problematica del bene giuridico e la scelta delle sanzioni*, in *Dei delitti e delle pene*, 1984, 114; P. NUVOLONE, *Le leggi penali e la Costituzione*, Milan, 1953; M. SPASARI, *Diritto penale e Costituzione,* Milan, 1966, M. ROMANO, *"Meritvolezza di pena", "bisogno di pena" e teoria del reato*, in *Riv. it. dir. e proc. pen.*, 1/1992, 39.

[130] F. PALAZZO-F. VIGANÒ, *Criminal Law. A conversation,* cit., 222.

[131] See most recently the collection of writings by G. INSOLERA, *Decline and Fall of Liberal Criminal Law*, cit. On this point, see also: E. DOLCINI, *La pena ai tempi del diritto penale illiberale,* cit.; M. DONINI, *Perché il garantismo non rende liberali la pena e il "diritto" penale, ma solo legittimi. Reflections in the margin of the Manifesto of liberal criminal law of the Union of Italian Criminal Chambers*, in *Cass. pen.*, 9-2019, 3132 ff. On this point, see also the volume AA.VV. (edited by F. GIUNTA, R. GUERRINI, V. MAIELLO, S. SEMINARA, F. SGUBBI), *Criminal law and the liberal paradigm. Tensions and involutions in contemporary times*, Siena, 2020.

[132] Gaetano Insolera has recently observed how "*liberal ferment, as a thought far removed from salvific emancipatory promises and palingenetic finalisms, in our field provides the levees to oppose the easy subjugation of penal affliction to the pursuit of ideologies or demagogic shortcuts, which sometimes evoke the splendor of torture.*" G. INSOLERA, *Why we cannot but call ourselves liberals*, in *Ind. pen.*, 1-2017, 5.

[133] Thus F. SGUBBI, *Il diritto penale totale,* cit., 28. Opposite, albeit afferent to a common core of considerations, is the thesis argued by M. DONINI, *Il diritto penale come etica pubblica*, Modena, 2014.

the reflux of pre-Enlightenment elements, among which the presumption to protect the *truth* can certainly be included[134] , must be avoided.

Absolute concept, undoubtedly known from religion[135] and philosophy[136] , but now *filtered*, for centuries, into positive criminal law, which is not concerned with punishing *mendacity* per se, but with protecting specific goods or interests from individual forms of aggression that are carried out through deceptive conduct or by the formation of documents or deeds that are forged, altered, or in which falsehood is attested[137] .

It is a conception that is being challenged by the most recent philosophical debate, in which the existence of *alethic rights is* beginning to be discussed[138] , but which, in our view, must instead be reaffirmed, especially in the face of the twofold risk of constructing a form of *digital censorship* of unwelcome opinions and also punishing the dissemination--in good faith--of untrue news.

Indeed, it seems to us that the recourse to *absolute* concepts, as we have seen very dear to authoritarian systems, is hardly compatible with the postulates of an *open* society and more consistent with the vision of those who historically are its *enemies*[139] .

---

[134] See on this point the paper by O. DI GIOVINE, *About a recent debate on "truth" and "criminal law*," in *Criminalia*, 2014, 539 ff.

[135] See for all the volume by T. PADOVANI-V. VITIELLO, *Non dire falsa testimonianza,* Bologna, 2011, in which Tullio Padovani notes how the judicial dimension of the eighth commandment is not the only, nor, perhaps, the most important one. In fact, the scope of the commandment, according to a well-established hermeneutic, introduces the prohibition against any false assertion: the prohibition against lying. *Ibid*, 119-120.

[136] Although, as is well known, this is a concept about which philosophical dispute has been heated since the time of Socrates. This is certainly not the place to report the terms of such a wide-ranging debate; therefore, we refer to the reading of some essential texts, to which we have also turned for a general framing of the problem and its tradition in the history of thought: F. AGOSTINI, *Introduction to Truth,* Turin, 2011; S. CAPUTO, *Truth,* Roma-Bari, 2015; H-G. GADAMER, *Truth and Method,* Florence, 2019; D. MARCONI, *For Truth*, Turin, 2007; G. VATTIMO, *Farewell to Truth,* Milan, 2009;

[137] For all: T. PADOVANI, *Lie and criminal law*, Pisa, 2014. For an in-depth and acute examination of the topic, see also the recent monographic study by A. GALLUCCIO, *Punishing the Dangerous Word,* Milan, 2020.

[138] F. D'AGOSTINI-M. FERRARA, *Truth to power*, Turin, 2019.

[139] K. POPPER, *The open society and its enemies*, Rome, 1996.