LEX LOCALIS

Social Communications Media - From Deregulation to Re-regulation

Editors:
Katarzyna Chałubińska-Jentkiewicz
Monika Nowikowska


§
LEX
LOCALIS

Title: Social Communications Media - From Deregulation to Re-regulation

Editors: Katarzyna Chałubińska-Jentkiewicz, Ph.D., Associate Professor (Kozminski University, Cellege of Law), Monika Nowikowska, Ph.D., Assistant Professor (War Studies University in Warsaw, Faculty of Law and Administration)

Reviewers: Ksenia Kakareko, Ph.D., Associate Professor (Warsaw University)

Kataložni zapis o publikaciji (CIP) pripravili v Narodni in univerzitetni knjižnici v Ljubljani

COBISS.SI-ID 219629059
ISBN 978-961-7124-25-5 (PDF)

First published in 2024 by
Institute for Local Self-Government Maribor
Smetanova ulica 30, 2000 Maribor, Slovenia
www.lex-localis.press, info@lex-localis.press

For Publisher:
assoc. prof. dr. Boštjan Brezovnik, director

Price:
free copy

https://doi.org/10.4335/2024.2 ISBN 978-961-7124-25-5 (PDF)

Available online at http://www.lex-localis.press.

Social Communications Media – From Deregulation to Re-regulation

Editors:
Katarzyna Chałubińska-Jentkiewicz
Monika Nowikowska

Maribor 2024

Social Communications Media - From Deregulation to Re-regulation

KATARZYNA CHAŁUBIŃSKA-JENTKIEWICZ & MONIKA NOWIKOWSKA

Abstract The term cyberspace refers to the totality of phenomena in a parallel space, which constitutes a new field of human activity to which behaviours and solutions applied in the real world are transferred. Legislators at various levels - both international and national - are introducing new regulations. This has led to the obsolescence of the phenomenon of impunity for illegal activities online. Cyberspace is more malleable than reality in terms of adopting or creating patterns. Its susceptibility brings conveniences as well as entirely new challenges for the legislator. The convenience is the ease of introducing regulations adequate to those in force in the real world, but the regulations so established are often met with obstruction or simple ignorance on the part of users of the ICT network, in particular due to the lack of instruments for redress or prosecution of crime. In the case of behaviour related to the functioning of cyberspace, also due to its global nature, such a relationship does not seem obvious. This is because activities in virtual space are characterized by their own specific culture of behaviour of its users - the virtual community. It should therefore be assumed that the new phenomenon of security required in the context of the functioning of ICT networks creates the need to take into account situations that may not be reflected in the world outside cyberspace.

Users of digital services are no longer just a passive party in the content delivery process, but have become active participants and are now both a source and a receiver of content in the digital ecosystem. Indeed, in terms of design, business model and optimization, information society services are based entirely on the concept of the dual role of their users.

CORRESPONDENCE ADDRESS: Katarzyna Chałubińska-Jentkiewicz, Ph.D., Associate Professor, Kozminski University, Cellege of Law, ul. Jagellonska 57/59, 03-301 Warszawa, Poland, e-mail: kasiachalubinska@gmail.com, ORCID: 0000-0003-0188-5704. Monika Nowikowska, Ph.D., Assistant Professor, War Studies University in Warsaw, Faculty of Law and Administration, Aleja Generała Antoniego Chruściela „Montera" 103, 00-910 Warszawa, Poland, e-mail: m.nowikowska@akademia.mil.pl, ORCID: 0000-0001-5166-8375.

https://doi.org/10.4335/2024.2 ISBN 978-961-7124-25-5 (PDF)
Available online at http://www.lex-localis.press.

The publication addresses the following issues: Security Risks and Public Risk Perception Associated with Digital Media, Disinformation and Cyberterrorism in Light of the Standards of the Council of Europe, Regulatory Dilemmas Around Social Media, Information, Disinformation, Cybersecurity, Content Blocking in Light of the Polish Broadcasting Act and the Digital Services Act (DSA) – Comments on the Mutual Relationship of the Acts, Disinformation in the Regulations of Selected Countries. The shift in thinking about the rights of web users is indicative of a change in the roles of the various actors that make up the current digital environment, the digital media. There seems to be a need for new regulation of social media. Recognition of the lack of such regulation at EU level provides a rationale for national regulation. A regulatory minimum is justified.

Keywords: • cybesecurity • cyberspace • digital content • disinformation • internet • new media • social media

Table of Contents

Introduction

Digitisation as a technological, but also social process has contributed significantly to the development of new technologies and, consequently, the so-called new media. New technologies are the subject of discussion and ultimately regulation, in terms of legislative action. The development of new technologies, as well as the associated processes of social change, require new regulatory approaches. It should be pointed out that the processes of convergence of digital media with traditional media have contributed to a particular type of conflict in the area of arrangements as to the scope and level of new regulation, both in the cultural field, where most issues concern new media and new technologies, such as the protection of intellectual property, the right to privacy, the protection of children and young people, and in the economic field: the rationing of the media market, the responsibility of digital service providers. New models of information management are emerging, including online. Changing communication technologies have fundamentally changed the rules for individuals and entire communities. An open and free cyberspace allows for the exchange of cultures and experiences between states, communities and citizens, enabling interaction and the exchange of information and, consequently, knowledge, experiences and technology. The ideological basis supporting this exchange is freedom of speech, freedom of communication. Digital reality allows public tasks to be carried out in a new social dimension.

The texts included in the present study are the result of the Authors' research on issues of scientific interest. The selection of issues in the publication is related to widely understood notions of media, communication, new technologies. The authors' aim was to bring the readers closer to the most important issues related to the functioning of new media and emerging problems.

In the article „Security Risks and Public Risk Perception Associated with Digital Media", Katarzyna Chałubińska-Jentkiewicz presents the impact of social media on the public sphere. Addiction to social media has been an ever-growing concern as the number of users and smartphone owners continues to rise, especially among young people. Yet, it is not only millennials who are exposed to Internet-related threats. Older people also spend more and more time online, including on social media. Research shows that social media abuse reduces psychological well-being and satisfaction with one's appearance. To some extent, it also influences the way we see the world. Prolonged exposure to social media can also lead to eating disorders, sleeping problems, and even fits of aggression. This article is the result of the author's research, as part of the internship she did in Italy at the Università degli Studi di Udine, Dipartimento di Lingue e Letterature, Comunicazione, Formazione e Società in 2022.

In the article „Disinformation and Cyberterrorism in Light of the Standards of the Council of Europe", Katarzyna Chałubińska-Jentkiewicz and Monika Nowikowska presented the

problem of disinformation in new media. Disinformation constitutes a serious security threat for contemporary democratic societies - states, international organisations, and individuals. It should be stressed that this phenomenon is becoming one of the most significant and complex challenges of the 21st century.

Jędrzej Skrzypczak, in his article „Regulatory Dilemmas around Social Media", analyses whether there is a need or necessity to regulate how social media operates. If such a need or necessity indeed exists, it would be warranted to consider the methods „hard" regulations or self-regulatory solutions, with which to fulfil them and the levels national, regional or international at which they should be implemented, to make sure the solutions are effective in the complex social-media environment. The analysis was primarily based on comparative, inductive and deductive methods, and on legal exegesis.

In the article „Information, Disinformation, Cybersecurity", Katarzyna Chałubińska-Jentkiewicz and Monika Nowikowska discuss new phenomena related to mass communication. A new phenomenon changing existing communication process rules is the so-called post-truth. One of the most significant global crises of our time, involving the spheres of political, social, and cultural relations and, later, the scope of mass communication, has been named post-truth. In 2016, the editors of the Oxford Dictionary declared post-truth the „word of the year". Such interest in the neologism is understandable, given the phenomenon which this word denotes. This article is the result of the author's research as part of her internship in Italy at the Università degli Studi di Udine, Dipartimento di Lingue e Letterature, Comunicazione, Formazione e Società in 2022.

Grzegorz Tylec in the article „Content Blocking in Light of the Polish Broadcasting Act (BA) and the Digital Services Act (DSA) - Comments on the Mutual Relationship of the Acts" analyses the issue of content blocking on the Internet. The comparison shows that although the Audiovisual Media Services Directive and, with it, the BA constitute lex specialis to the DSA, this legal act will largely shape how modern online media functions and will do so on the same basis for all EU countries. It can be seen, from the comparison, that the DSA, unlike the BA, will apply to the operation of social media and, in addition, it will also cover the activities of platforms, regardless of whether their providers have the status of business entities. It should be assumed that, even though, formally, the DSA constitutes lex generalis to the BA, its provisions will be applied alongside or in parallel with the procedures envisaged in the BA. This is because it is difficult to argue that the applied procedures provided for in the BA would preclude the actions provided for in the DSA.

Katarzyna Chałubińska-Jentkiewicz in the article „Disinformation in the Regulations of Selected Countries" analyses the issue in question on the grounds of various regulations, analysing, inter alia, the solutions of the Republic of China, the Russian Federation, Spain, Sweden, Norway, Canada or Israel.

Chapter I

Security Risks and Public Risk Perception Associated with Digital Media

KATARZYNA CHAŁUBIŃSKA-JENTKIEWICZ

Abstract The media system is the simplest reflection of the social and political situation in a country. The media are also important for state policy, for understanding the value of the public interest. This is especially the case when, when the media are analyzed in the context of their paternalistic role, i.e. their public mission in terms of educating the public, active in public life, according to a sense of axiology and national identity, etc. This is all the more important because in the media market, significant even revolutionary changes are taking place, leading to its liberalization. This is creating the conditions for the development of an alternative media system based on the principles of competition and the provision of a so-called digital service. At the same time, media rules are being unified on a European scale, which particularly concerns new technological conditions and the protection of market rules. In this area, it is obvious and desirable to create completely new solutions enabling the exchange of experience and the preservation of basic requirements defined at EU level. Therefore, the creation of a strengthened organizational system is needed, but without interfering in the regulatory area of the EU Member States. The digital media environment is subject to obvious changes, but it is still the state that plays an important role as a regulator of media reality. At the core of the functioning of a democratic state under the rule of law is freedom of expression. Therefore, on the one hand, the media are a check on the activities of the authorities, but on the other hand, they are also subject to such supervision. It is important to emphasize the position that the specific role the media play in the state and society needs to be assessed from the point of view of the overriding good, which is the public interest in national terms.

Keywords: • digital content • social media • public interest • service users

CORRESPONDENCE ADDRESS: Katarzyna Chałubińska-Jentkiewicz, Ph.D., Associate Professor, Kozminski University, College of Law, Jagiellonska 57/59, 03-301 Warszawa, Poland, e-mail: kasiachalubinska@gmail.com, ORCID: 0000-0003-0188-5704.

https://doi.org/10.4335/2024.2.1 ISBN 978-961-7124-25-5 (PDF)
Available online at http://www.lex-localis.press.

If men were angels, no government would be necessary. If angels were to govern men, neither external nor internal controls on government would be necessary. In framing a government which is to be administered by men over men, the great difficulty lies in this: you must first enable the government to control the governed and, in the next place, oblige it to control itself.

J. Madison

1 Introductory remarks

Answers to the questions concerning the methods of applying the law, as an instrument to influence the market and public governance, should be sought for regulation purposes and for redefining public interest objectives. The primary purpose of digitisation is to secure the development of society in the digital consumption era. Due to the passage of time and the development of new transfer techniques, the digital resources meeting the adopted standards related to their legal sharing are dramatically diminishing. Digital content constitutes a source of knowledge, inspiration, and skills for present and future generations. The underlying public tasks related to digitisation include organising digital resources and managing them modernly. The present and future duties of entities and participants in the market for digital services in the digital world need to be normatively established. This requires infrastructure owners to cooperate with owners of digital resources and the respective public institutions of e-administration.

Another significant objective is to provide Internet access to digital content. Since Internet technologies have been popularised, and more and more materials are being shared via the Internet, it is becoming necessary to determine which priorities should be adopted to lay down the principles of access to resources after the digitisation process. Should it be a common public service? Or should access be limited to information security (classified information, personal data, copyright, etc.) or to private or economic interest?

The questions relating to public governance in cyberspace also apply to the matter of implementing the directives on the reuse of information from the public sector. Encouraging people to reuse digital content constituting the public domain is obvious. However, the reuse of digital resources, e.g., by the private sector, raises many problematic issues, such as intellectual property rights, data protection, policy concerning the collection of fees, and the market competitive balance between public and private services.

The management of digital content-sharing processes changes due to the increasing commitment of various social groups to the functioning of a modern information society. One should bear in mind that the contemporary digital resources shared with the use of digital media, which were often developed from the traditional ones, must be a reliable and legal source of knowledge. At the level of the contemporary development of the digital society, it is necessary to lay down the rules for managing digital resources at all

stages of digital content circulation. Just as public institutions must participate in the management of market processes at the level of both the government and the local-government administration sectors, it seems that they should do so in both the technological aspect and through the development of new strategies and planning. The last matter to be resolved is the significance of the legal principle. According to R. Cooter and T. Ulen, the economic analysis of law can be divided into positive (dealing with evaluating the effects of particular regulations regarding their economic efficiency) and normative (providing broadly-construed recommendations and postulates regarding legislative activity). The case described by R. Cooter and T. Ulen, in their book titled *Law and Economics*, can serve as an exemplary solution (Cooter, Ulen 2011: 166). Referring to the problem of the future protection of copyright in cyberspace, the authors provide several solutions. It might be collective management, the so-called celestial jukebox, under which every digital information user will pay royalties to a central clearing house managing the copyright. Then, copyright will become the dominant law of the digital age. Another solution is "digital libertarianism", in which technical protection through cheap encrypting will be more efficient than the legal protection of intellectual property, and copyright law will die out because technology will make the law unnecessary. Perhaps this is the fate of many other regulations in the modern world, as we still do not know whether new laws respond to new mechanisms or contrariwise (Cooter, Ulen, 2011: 166–167).

Driven by digital technology advancements, the functioning of the European media sector in the digital age is generating increased consumer demand and globalisation. These processes are posing new challenges for the regulators. Digitisation is defined as the conversion of a signal (e.g., processed sound, images, and data) from analogue into digital form through analogue-digital processing. Also, it means converting analogue format into digital (binary) format, which can be stored in computer memory. The term *digitisation* stems from the word *digit*, which originated from the Latin *digitus* (finger, toe, counting fingers). On 3 October 1997, the European Commission issued the already inapplicable Communication No. 623 – the Green Paper on the convergence of the telecommunications, media and information technology sectors, and the implications for regulation (the Green Paper on the convergence of the telecommunications, media and information technology sectors, and the implications for regulation towards an information society approach, COM(97) 623 (non-applicable version).

In the Green Paper, the Commission highlighted that computer technology played a key role in content creation and broadcasting. The digital media field has since, however, undergone enormous change, with content – and, more specifically, digital content – being now created by all users in the digital world.

Due to its widespread accessibility and social impact, the Internet creates the most extensive opportunities for everyone to participate in political and cultural life by creating and sharing digital content. The development of the information society is accompanied

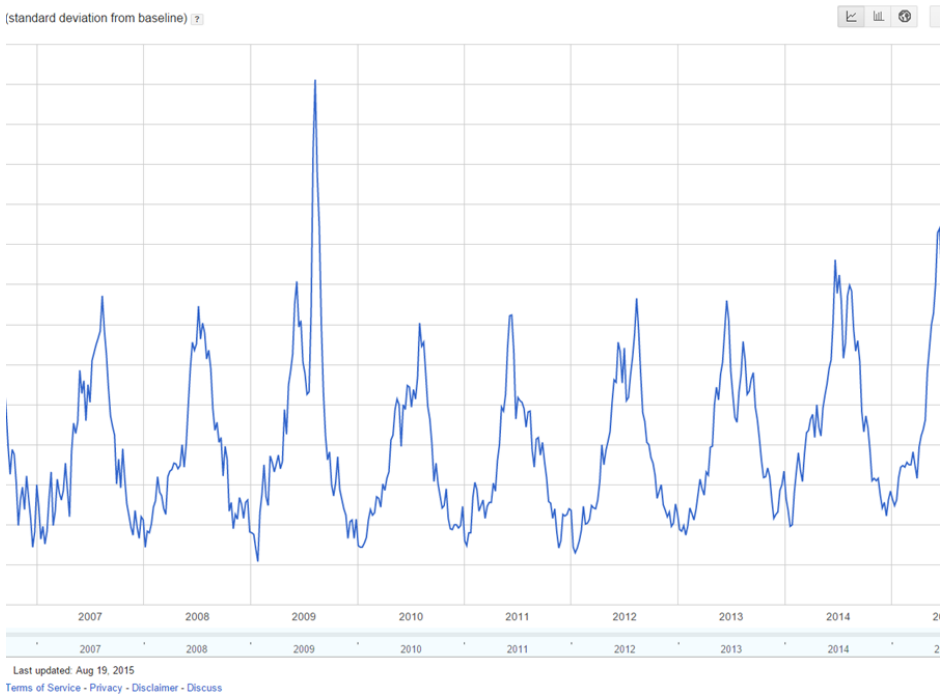
by civilisational as well as economic and cultural development. Mass communication is becoming a source of knowledge about the world, driving the emergence of a mass and global society. Each day, the Google search engine responds to three billion queries, and all of them are archived. Big data poses a challenge to outer interactions with the world (a free online service of Google developed to translate texts, text files, websites, speech, pictures, and videos into 100 languages in real-time. While a text is entered, the service translates it in real-time. If a single word is put in, Google Translate works like a dictionary, usually providing from several to a dozen-odd suggested meanings. Google Translate is used by 200 million people daily. Since September 2016, Google Translate has had a new engine, GNMT, Google Neural Machine Translation – based on recurrent neuronal networks. The program now translates whole sentences (whereas earlier it translated single words), significantly improving the quality of the target-language text. It has supported the Polish language since March 2017).

GNMT improves translation quality because the system learns from millions of available examples. By using this broader context, it can find the most accurate translation. It then processes and adapts it to more "human" speech with the correct grammar. By using all these data, we can identify links and details which would otherwise be lost in the sea of information. The most interesting information is that which deviates from the norm, and the only way to identify it is by comparing vast numbers of transactions. For example, the unauthorised use of credit cards is detected by searching for anomalies. As the amounts of data grow, so do inaccuracies since large datasets always contain incorrect figures and distorted information. Yet, big data compensates for this lack of order.

Society benefits from big data not because of the faster processors or improved algorithms but because of the larger quantities of data. It is not causality but correlations that will be searched for. Around seven billion stocks trade hands on the US stock exchange daily. Approximately two-thirds of all these transactions are initiated by computer algorithms, which process vast amounts of data to bring profit at an acceptable risk. Facebook processes 10 million new pictures every hour, and 800 million YouTube users upload an hour of new videos every second. Each year, the volume of messages on Twitter is growing by about 200%. Big data relies on prediction. Amazon can recommend books to users, Google can display the requested website, and Facebook knows what its users like, while LinkedIn can guess whom users know or might know. Twitter, LinkedIn, and Facebook create "sociograms" of their users to identify their preferences (Microsoft acquired Farecast in 2008. It was an online booking portal publishing predictions on the best times to buy airline tickets. Farecast was founded by the American scientist Oren Etizioni in 20003. In 2007, it recorded more than 175 billion views of airline tickets. Farecast's data monitoring team used airline ticket price observations to develop algorithms predicting future price movements. In May 2008, Microsoft integrated Farecast's website with the Live Search engine to create the Live Search Farecast program, registering it in June 2009 as Bring Travel as part of the work on developing new search mechanics).

Thanks to large datasets, decisions can be made not by humans but by machines. There is a growing awareness in Poland that efforts should be made to support digital integration. The problem, however, lies in the lack of a systemic approach to digitisation and the lack of a coordinated approach to digitisation initiatives, including regulatory ones, causing the duplication of efforts and the underperformance of measures. The media sector has a key role to play in developing European citizenship, as it is one of the core means of communicating the common, fundamental social and cultural values of the Union to European communities, particularly young people. Digitisation aims to secure these in the form of high-quality digital copies and also to provide users with the broadest possible access to national heritage resources by creating online archives, library resources, and digital repositories. Digitisation and network technologies are essential drivers of economic and social development. Importantly, these phenomena are not only the domain of the public sector. Rather, it is private resources which largely constitute the most valuable sources of knowledge about their owners, holders, and users. And this, in the current age of profiling for marketing and other purposes, represents the most valuable marketable good. Digital content protection objectives can be achieved only by ensuring that digital initiatives, thus far largely dispersed, are coordinated in terms of creating resources, providing them with long-term protection, and establishing fair and transparent conditions for their sharing. Social media are essential vehicles for sharing content on the Internet.

Figure 1: Google Flu Trends Data



Google Flu Trends is an online service designed to estimate the number of flu infections in more than 25 countries. By linking Google search queries, Google wanted to make accurate projections about the spread of influenza. The project was launched in 2008 by Google.org to prevent the spread of flu epidemics. Google Flu Trends is no longer publishing current estimates. Historical data are still available, while current information is provided only for research purposes.

2 The impact of social media on the public sphere

Addiction to social media has been an ever-growing concern as the number of users and smartphone owners continues to rise, especially among young people. Yet, it is not only millennials who are exposed to Internet-related threats. Older people also spend more and more time online, including on social media.

Research shows that social media abuse reduces psychological well-being and satisfaction with one's appearance. To some extent, it also influences the way we see the world. Prolonged exposure to social media can also lead to eating disorders, sleeping

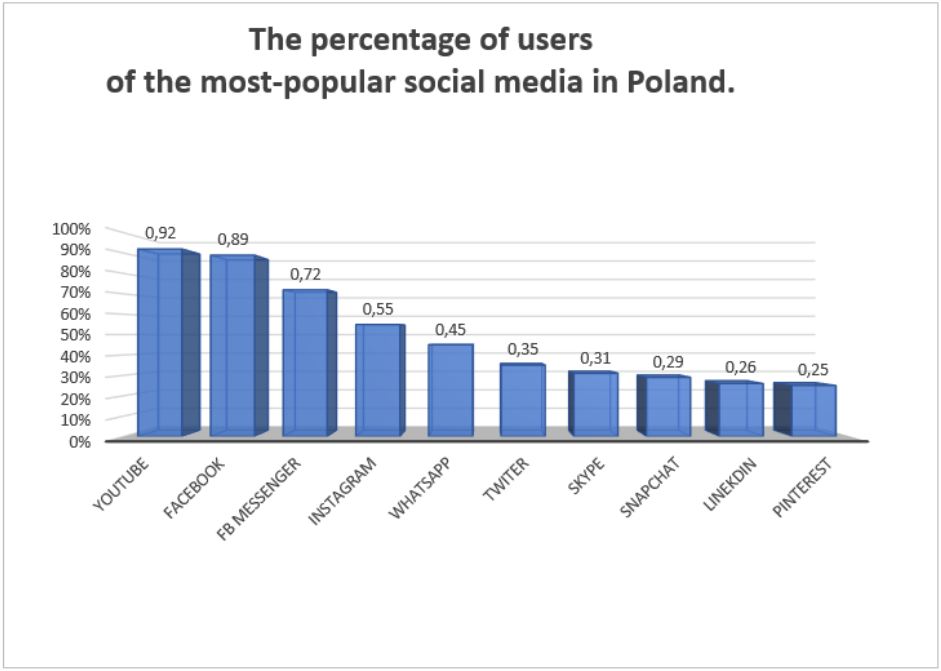
problems, and even fits of aggression. Fortunately, it is increasingly the subject of social and scientific debate, raising awareness about these critical and direct threats to Internet users. According to S. Galloway, "It's easy to be sceptical about Facebook, especially with all of the self-promotion, fake news, and groupthink spread on the platform. But it's also hard to deny it nurtures relationships, even love. And there is evidence that these connections make us happier" (Galloway, 2018:130). This is clearly the case with my retired father, who has now connected with his schoolmates with whom he had had no contact throughout his professional and private lives. As well as reconnecting with old acquaintances, Facebook allowed him to discover new interests, and even the theft of his identity and the blocking of his account did not stop him from surfing Facebook.

However, apart from the defined threat involving additions from the Internet, there are other issues with potentially much broader implications.

Although browsing social services is widely abused (mainly out of boredom), it can also be used more consciously, including for business purposes. Social media are excellent tools to find varied information and to get consumer feedback. This is well exemplified by Facebook groups, which focus on a single, often niche, topic or issue. By visiting these groups, we can get lots of essential information on a subject of interest to us or talk to people who seem to have similar interests (www.whysosocial.pl).

In addition to functional, strictly social advantages afforded by such services, social media are currently used as vehicles for influencing public opinion. It is not uncommon for election campaigns to rely substantially on the Internet to manipulate facts to undermine other candidates. This has become commonplace now, with social media being sources of information with questionable reliability. Furthermore, users are inundated with product advertising campaigns and sponsored messages designed to match their behavioural patterns, habits, and individual preferences. All these elements of being part of the Internet community generate a sense of chaos and uncertainty about the rules governing the digital market.

Chart 1: The most popular social media in Poland



Source: Based on the #Digital2020 report,https://mobirank.pl/2020/02/23/digital-mobile-i-social-media-w-polsce-w-styczniu-2020-roku/ (10/07/2020).

According to Digital2020 (www.datareportal.com), as of January 2020, there were more than 19 million social media users in Poland. YouTube, the video-sharing platform operator, was the leading social media site in Poland, with 92% of Internet users in Poland accessing the resources of this online service. YouTube was followed by Facebook, with 89% of Internet users. At 72%, the third place was taken by Messenger, followed by Instagram (55%) and WhatsApp (45%). Those further down the chart are clearly less popular, with fewer than 50% of Internet users reporting their presence on these services.

Social media are now undoubtedly the largest source of information and the leading space for processing digital content. On the one hand, such media expand access to sources of information and, on the other, the globalisation and availability of digital resources of various kinds and values cause users to be closed in "information bubbles", in which they function within a restricted circle of personalised information. This creates a kind of separation from other content outside the scope of the user's interests. The reason is that, as marketing and advertising algorithms inundate users with advertisements of the product they are looking for, the algorithms used in content searching match the content

to the interests and the users' specific needs. These activities are mainly self-regulated through the terms of use of individual social media. For instance, in Facebook's Terms of Service, the need for personalisation is expressed as a suggestion (www.facebook.com/legal/terms). It should be added that it is enough to talk about or express any interest using a smartphone, or "near a smartphone", for the mobile device to singly decipher and classify user interests as part of profiling. Nearly every movement, voice, and sign of interest plays a role. Not to mention geolocation-based analyses. A crucial task of social media is to attract the audience's attention. YouTube is an example of this, as it introduced a simple, and yet effective, method of attracting audiences merely by playing videos with related subjects one after another. Netflix, a provider of audiovisual on-demand services, adopted a similar approach of suggesting related content after watching a video or playing another episode in a series. User attention is the primary objective of Snapchat, the major messaging app used by teenagers. It should be noted, however, that whilst it is more of a communication aid for adults, for underage users, the app constitutes a *centre* for communication. Snapchat has introduced "Snapstreaks" – an update which tracks the number of days during which two users exchange "snaps" daily. Snaps are not necessarily regular conversations but pictures of a street, wall, or other objects without specific information. Still, data obtained from content shared in this way are used for analysis and drawing conclusions. Data are employed to analyse information about users and their environments. "A privacy advocate's nightmare is a marketer's nirvana. The open nature of Facebook, coupled with the younger generation's belief that 'to be is to share', has resulted in a data set and targeting tools that make grocery store scanners, focus groups, panels, and surveys look like a cross between smoke signals and semaphore (...). When you have the Facebook app open on your phone in the United States, Facebook is listening… and analysing" (Galloway, 2018:131).

With such vast amounts of data regarding users, which are available to them, media companies can not only reap substantial profits but also, in a sense, create a digital reality. As an example – consider the "PizzaGate" conspiracy theory, according to which officials affiliated with the former US presidential couple and members of the Democratic Party, including members of the Washington elite, were involved in a paedophilia ring. The central premise on which online investigators based their claims was that a paedophilia ring had an affiliation with a pizzeria. "Cheese pizza" is one of the English expressions used to refer to child pornography, its abbreviation "CP" also being understood as "Child Porn". The scandal erupted in early November 2016 when WikiLeaks published another batch of emails by John Podesta, Hillary Clinton's campaign manager. Unidentified perpetrators hacked into Podesta's account. Then, amidst the tension surrounding the upcoming elections, they stole data on email correspondence. However, a new batch of the campaign manager's emails was published by a user of the popular Internet forum, 4chan. The post implied that, this time, the correspondence contained major revelations, purporting that Podesta was a member of a paedophilia ring. This led to a proliferation of analyses on Reddit, and the "PizzaGate" subreddit was created for all topics around "PizzaGate", bringing together around 22,000 followers, including 1,500 active users.

This scandal was part of a broader trend of fake news, often craftily fabricated facts and lies. The very place these discussions were held was suspicious, with 4chan and Reddit being notorious for Internet memes. People came to believe that the owner of the Comet Ping Pong pizzeria was involved in child trafficking and molestation. They also claimed that Hillary Clinton and John Podesta were part of the conspiracy. As Comet Ping Pong began receiving telephone threats, fiction became reality. These attacks were initially only unsavoury jokes about "a special pizza" or other absurd names used to allude to paedophilia. Eventually, though, they became actual threats: "We'll kill you all", "We know where you live, you better kill yourself". One of the threats said "I'll go the restaurant with a rifle and kill you all, I'll rip your guts out and watch you die like an animal". The culminating point came when Edgar Maddison Welch from North Carolina arrived at the pizzeria, believing that children were being imprisoned and raped in the establishment's basement. A father of two, Edgar was convinced that he was on a rescue mission. One Sunday afternoon, when the pizzeria was full of families with children, the man entered the restaurant. He was armed and screamed at the people, telling them to leave the building. Then, he proceeded to search it. After rummaging through the entire establishment, he found only a locked door, which he shot open. The door led to the backrooms. Here, he saw a small sanitary room instead of a basement with children imprisoned. Edgar Maddison Welch was ready to go to jail or die because of false, misleading information (Rossi, 2020:498). The PizzaGate story demonstrates the magnitude of the impact that content created on social media can have on Internet users.

Hence, it should be assumed that the authenticity of digital content online depends not on accuracy, truthfulness or reliability. It is rather a matter of user reactions, such as 'likes' and reposting. These also create digital content which defines social preferences. In addition, such behaviour is used for political and criminal purposes. Generally speaking, it can be assumed that social media initially became a self-sufficient form of communication for organising protests and enabling activists and citizens to communicate their specific needs (e.g., through tweets calling for blood donors). Mainstream journalists realised that social media users had access to information or voided materials to which they had no access. This was due, for instance, to bans imposed by media organisations or Internet and telecommunication network blockades. Meanwhile, social media users were checking content for legitimacy and accuracy, aware that social media are not regulated in this respect (Bebawi, Bossio, 2014:135).

3 The responsibilities of digital content providers

The business of digital content providers consists of making content available through information and communication systems. This category is highly diverse. It includes not only specialised institutions but also end users. The latter group is particularly active due to the growing popularity of user-generated sites (or user-generated content – UGC).

Due to the active form of online operations, content providers seem to bear direct liability for any breaches caused by such operations. In Poland's legal system, content providers are directly liable for infringements of third-party rights. As noted by J. Barta and R. Markiewicz, controversies arose around attempts at qualifying the act of making works available on computer networks. Ultimately, this was qualified as a new field of exploitation, i.e., making a work available in a way that it could be accessed by anyone at any time and place they choose (Barta, Markiewicz, 2001: 228). This issue was highly relevant for ICT networks, whose function was based on interactivity. As a result of digital processes, users can modify and share content without problems. The concept emerged of *sui generis* protection for the rights of online content producers or providers. It was discussed at the Association Littéraire et Artistique Internationale (ALAI) congress in 1996, with attempts to formulate a construct allowing producers to claim protection against third parties. Among others, consideration was given to affording them the status of moral rights or quasi-moral rights, with the caveat that they might not have limited the moral rights of content creators (Dietz, 1997; as cited in Gęsicka, 2014:290). According to J. Barta and R. Markiewicz, the construction of these rights is similar not to moral rights but to the economic rights vested in authors (Barta, Markiewicz, 2001: 228). It was this core objective, primarily economic, that these entities had in mind, bringing these rights closer to related rights.

As regards other infringements, direct liability was also assigned to content providers. Therefore, they were no longer exempt from liability, which used to be restricted to the suppliers of electronic services. Technological changes influenced the scope of liability for illegal acts in cyberspace. Also, new rules on the limitation of this liability were introduced. In European law, the liability of online service providers is regulated by Directive 2000/31/EC, which lays down the rules governing the liability of digital content intermediaries. The Directive contains provisions on the most popular online services, i.e., mere conduit, caching and hosting. It should be noted that the European regulation follows the horizontal model. This means that the exemptions it provides apply to any legal liability, including civil, criminal, and administrative liability. The Electronic Commerce Directive lays down the rules for excluding liability at the maximum level. Consequently, individual Member States may decide to impose less strict solutions. Similar liability rules are laid down in the draft Regulation of the European Parliament and of the Council on a Single Market for Digital Services (the Digital Services Act) and amending Directive 2000/31/EC COM/2020/825 final.

The provisions of the Electronic Commerce Directive were implemented into Polish law by Articles 12–15 of the APSEM. Under Article 12 of this Act, relating to the mere-conduit service, "no liability for the provided information shall be assigned to those transmitting data who 1) have not initiated the transmission; 2) have not chosen the recipient of the data; and 3) do not remove or modify the data transmitted. The exclusion of liability referred to in paragraph 1 extends also to the automated short-term indirect storage of the transmitted data, provided that this is required only to complete the

transmission and that data are not stored longer than necessary in normal circumstances for completing the transmission (caching)" (Article 1(2) of the APSEM).

Caching – etymologically deriving from the French word *cacher*, meaning to hide or conceal – is an automated process of creating temporary copies of digital data to allow greater data accessibility for more frequent use. Caching is permissible as an exception to the right to reproduce a work under Article 5(1) of Directive 2001/29/EC. This provision establishes the rule according to which specific acts of temporary reproduction, which are transient or incidental reproductions, form an integral and essential part of a technological process and are carried out for the sole purpose of enabling either efficient transmission in a network between third parties by an intermediary, or the lawful use of a work or other subject-matter, to be made.

In the case of the caching service, the waiving of liability for storing data applies to entities which transmit such data and ensure their automated and short-term indirect provision to help other entities re-access them on request, but which 1) do not remove or modify such data; 2) use recognised and customary IT techniques defining the technical parameters of data access and updating; and 3) do not interrupt the use of recognised and customary IT techniques to collect information about the use of the data gathered (Article 13(1) of the APSEM). Hence, respecting the integrity of the stored data is a prerequisite for avoiding legal liability. Under Article 13(2) of the APSEM, "no liability for stored data shall be assigned to those who, subject to the conditions referred to in paragraph 1, immediately remove, or prevent access to, such data, on becoming aware that the data have been removed from the original source of transmission, or access to them has been prevented, or when a court or other competent authority ordered that such data be removed, or access to them be prevented".

Article 14 of Directive 2000/31/EC should be interpreted as implying that the rule laid down by it applies to the provider of an Internet referencing service if such a provider supplies such services without playing an active role, which could provide them with knowledge about or control over the information stored. If a service provider does not play such a role, it may not be held liable for the contents of the information stored at an advertiser's request unless immediate measures were not taken to remove or prevent access to such information once the service provider became aware of the illegal nature of such information or the advertiser's business.

Thus, the scope of liability of website providers is influenced by the type of their services. This also applies to various fields of content regulation. Assuming we are dealing with a website which meets all the definitional requirements of the press, the activities of such a provider are subject to press registration under the procedure set out in the Press Law Act (Article 20 of the Press Law Act – the regional court register; providers of audiovisual media services are exempt from this obligation under Article 24 of the Press Law Act, consolidated text, Journal of Laws of 2018, item 1914). In this case, liability rests with

the publisher and the editor-in-chief. As regards audiovisual media services, the applicable regulation is set out in the Broadcasting Act (consolidated text, Journal of Laws of 2022, item 1722); Article 41 of the BA – KRRiT (National Broadcasting Council) register or licence, depending on the type of dissemination – the register for programmes disseminated only on communication and information systems; licences for dissemination on communication and information systems and the broadcasting of programmes on operator systems) (Article 33 of the BA – KRRiT licence). Broadcasting liability will also apply here, and each case will involve editorial responsibility. It is slightly different with liability for making on-demand audiovisual content available. While business restrictions do not apply here, the rules of liability do because of extending the BA rules to providers of on-demand audiovisual media services. Under Directive 2010/13/EU, the extension of regulations to non-linear media services was meant to be restricted to on-demand audiovisual media services, excluding on-demand audio media services.

The extension included:

1) a stipulation that on-demand audiovisual media services also serve the functions assigned to broadly-defined radio and television broadcasting and, as such, they constitute parts thereof;
2) consistently replacing the word *programme* with the phrase *media service* (or, optionally, adding on-demand audiovisual media services next to programmes), and the word *broadcaster* with the phrase *media service provider*. Where the regulation applies to both programmes and on-demand audiovisual media services, or providers of both types of service, this extends to:
   a) freedom of reception (Article 1(2) of the BA),
   b) jurisdiction (Article 1a of the BA),
   c) the powers of the National Broadcasting Council (KRRiT) (Articles 6 and 10(2–4)of the BA),
   d) legal liability (Article 53(1) and Article 53a of the BA);
3) providing a reference to the appropriate application of some of the basic programme-related requirements to on-demand audiovisual media services in respect of the protection of minors and the promotion of European broadcasts: the freedom of the broadcaster to shape the on-demand media service (Article 47a and 47b of the BA); the identification obligations of the on-demand media service provider (Article 47c of the BA); the obligation of the easy recognisability of commercial communications (Article 47k of the BA); the general rules applicable to commercial communications: advertisements, sponsoring, telesales, product placement (Article 47k of the BA); the prohibition of discrimination and incitement to hate (Article 47h of the BA); the obligation to ensure that the services are available to people with visual and hearing disabilities (Article 47g of the BA); requiring that on-demand media service providers follow the rule of editorial business, within the meaning of the Press Law Act (Article 47a of the BA); the obligation to record broadcasts and

advertisements (Article 47i of the BA); and the obligation to issue reports to KRRiT (Article 47j of the BA).

Concerning on-demand services, Directive 2010/13/EU already provides for the same high level of protection for many elements, including service provider identification, the total prohibition of incitement to hatred, and quality standards applicable to audiovisual commercial communication. Notably, under Article 3(1) of the APSEM, the provisions of this Act do not apply to the dissemination or distribution of radio or television programmes, or any related text communications. However, this regulation relates exclusively to traditional radio and television communications, which are not provided on demand. Conversely, new media services, which fall under the Broadcasting Act following the implementation of Directive 2010/13/EU, fulfil the requirements of the Act on the Provision of Services by Electronic Means, and should be governed by them, including by provisions on the limitation of liability.

Another piece of EU legislation governing liability for online content-sharing is Directive 2001/29/EC, which introduces limitations on liability for copyright infringement. Article 5(5) of Directive 2001/29/EC allows the exemptions related to illegal use, provided for in Article 5(1–4), including the exemption for making copies for private use, as referred to in Article 5(2b) of this Directive, subject to the following three conditions: 1) such an exemption may be applied in certain exceptional cases only; 2) it does not conflict with a normal exploitation of the work; and 3) it does not unreasonably prejudice the legitimate interests of the rightsholder. As explained by Recital 44 of Directive 2001/29/EC, these three conditions correspond to the international obligations of Member States and the Union, and more specifically to the conditions applicable to all limitations of copyrights set out in Article 9(2) of the Berne Convention, more broadly known as the "three-step test," as reiterated in Article 13 of the TRIPS (the Agreement on Trade-Related Aspects of Intellectual Property Rights www.eur-lex.europa.eu) and in Article 10 of the WCT (the WIPO Copyright Treaty, Geneva 91996, www.eur-lex.europa.eu). The test will also apply to situations involving the use of digital content.

The examples provided above support the claim that, in each case, the same entity will be subject to different liability, depending on whether it is engaged in the service activities referred to in the APSEM, is a broadcaster or publisher, provides on-demand media services, or only provides a file-sharing platform. As a result of technological and economic convergence, the same entity can serve several different functions. Thus, its status – and, by extension, the scope of liability – is not definite. This calls for appropriate regulations providing that synchronisation is ensured at each stage of substantive legislative work. This is critical to establishing a cohesive regulatory framework which facilitates the development of the digital media sector while having due regard to the elementary principles of liability for disseminating digital content, in particular in the social media environment.

4 The rights of service users in the digital environment

Digital service users no longer play a passive role in the content communication process. Instead, they have become actively involved as both the sources and recipients of content in the digital ecosystem. Indeed, information society services base the entire design, business model, and optimisation of their services around the dual role of their users.

Regarding copyright protection, Internet users' activities have directly impacted the regulatory exemptions related to the digital non-commercial and proportionate use of quotations and extracts from copyright-protected works or other subject-matter by individual users. Under the Copyright Directive, subject to Article 13, Member States may provide an exception for content uploaded by users where such content is used for criticism, review, illustration, caricature, parody, or pastiche. Here, the question arises about the limits of such acceptable criticism, thus far primarily the domain of online journalistic activities. This change in the perception of Internet users' rights implies a shift in the roles played by various players in the existing digital environment and digital media.

It seems that attributing regulations directly to the need for a comprehensive remedy to issues around social media, and considering the lack thereof at the EU level as a legitimate reason for national regulations' going beyond implementing the Directive, is something of a simplification. Rather, definite guidelines should be followed on the scope of regulations, not necessarily involving comprehensive regulatory solutions anchored in national laws. Of course, it does not mean that this approach is wrong. However, it is more reasonable to rely on the regulatory minimum, given the need for arrangements regarding the use of digital content on European platforms. Recently, it has become necessary in this regulatory field to broaden the notion of a market for various types of content since its scope goes beyond the existing notion of a digital content market. This has been proposed in the draft Regulation of the European Parliament and of the Council on cross-border portability of online content services in the internal market (OJ L 168, 30/6/2017, pp. 1–11).

The document concerns the cross-border portability of online content services to which consumers have lawful access or content that they have purchased or rented online in their country of residence, and content they wish to continue to have access to when travelling within the EU. It also claims that the absence of or problems with the cross-border portability of online content services in the EU result from the licensing practices of copyright or related rightsholders and/or the commercial practices of service providers. Such services include websites which use works or other protected subject-matter only in an ancillary manner, such as graphical elements or music used as background, where the main purpose of such websites is, for example, the sale of goods. This means that practically every other use would be subject to the Digital Single Market regulations applicable to digital content, falling under framework regulation under this Directive.

The dichotomy of online services is recognised by the "Digital Single Market Strategy for Europe" (COM/2015/0192 final), which is built on three regulatory pillars: (1) providing better access to digital goods and services across Europe; 2) creating the right conditions and a high-level playing field for digital networks and innovative services to flourish; and 3) maximising the growth potential of the digital economy. It should be noted that the Union regulation concerns the regulation of the Digital Single Market, whilst the regulation of content (e.g., content protected by copyrights and related rights, and not only) should be considered as a matter of national-level solutions, having due regard to the specific needs of a given state, including in particular its national culture and cultural security.

An important factor in the context of users sharing their digital content online is the rule under which copyright-protected content is subject to licensing. In accordance with the Directive on Copyright in the Digital Single Market, the use of protected content by information society services providing automated image referencing is subject to obtaining a licence from rightsholders. Member States shall ensure that the information society service providers that automatically reproduce or refer to significant amounts of copyright-protected visual works, and make them available to the public for indexing and referencing purposes, conclude fair and balanced licensing agreements with any requesting rightsholders to ensure their fair remuneration. Such remuneration may be managed by the collective management organisation of the rightsholders involved, considerably facilitating the use of such content on online platforms.

Consideration should be given to issues concerning individual licences issued by competent, i.e., central authorities (followed by extended collective licences). The former determines state influence on how the digital content market functions but does not directly involve performing specific public tasks. While it seems debatable whether both legal solutions have the same strengths and weaknesses, they have the undeniable asset of affording users and owners of works considerable legal certainty, ensured by a single state authority which issues licences to use works on the Digital Single Market.

Another factor affecting the legal situation of Internet users is the processing of their personal data. A critical element in regulating Internet users' online activities is protecting their data. Under the GDPR (OJ L 119, 4.5.2016, pp. 1–88), data provided by data subjects are subject to specific rules. In particular, these include the right to data portability, which is limited to data "provided" by the data subject (rightsholder) to the data controller. In other words, it seems that user-provided data are the only information explicitly recognised as a "commodity", a kind of digital good owned by individuals.

On the one hand, this is the only dataset "portable" from one platform to another. On the other, it is the only type of (personal) data falling under a regulation which legitimises exchange, other than monetary, for supplying such digital content. Of course, "user-

generated" content does not always constitute personal data (as defined by Article 4 of the GDPR), just like not all user-provided personal data are valuable regarding intellectual property. Thus, a whole new field is opened for investigating content which is marketed in the digital economy but does not meet the definitional requirements set out by the regulations, currently accounting for various aspects of content sharing. This is particularly the case when user-provided private content is exchanged as a digital currency between data subjects (or users) and data controllers.

What should be stressed here is the importance of anonymity in processing personal data for commercial purposes. On the one hand, it is fundamental to support automated settings disabling personal data collection in the context of using online platform interfaces. On the other, such anonymity makes it harder to track Internet users who make illegal content available. This triggers a question of whether such protection creates a liability to pass entirely to digital service providers or whether the user bears some liability, too. "You don't have to share all data. But if you do, and data is sensitive, you should be able to do so in a manner where data can be trusted and protected. We want to give businesses and citizens the tools to stay in control of data. And to build trust that data is handled in line with European values and fundamental rights", said Margrethe Vestager, Executive Vice-President for A Europe Fit for the Digital Age. Her stance was supported by Commissioner for Internal Market Thierry Breton, "We are defining today a truly European approach to data sharing. Our new regulation will enable trust and facilitate the flow of data across sectors and Member States (...). With the ever-growing role of industrial data in our economy, Europe needs an open yet sovereign Single Market for data. (...) our regulation will help Europe become the world's number one data continent". The new Regulation will create the basis for a new European mode of data governance, in line with EU values and principles, such as personal data protection, consumer protection, and competition rules. This new approach proposes a model based on the neutrality and transparency of data intermediaries, which, as organisers of data sharing or pooling, may not deal in data on their own account (e.g., by selling them to another company or using them to develop their own products based on such data). Other legal solutions envisaged in the draft Regulation include measures to facilitate the reuse of certain data held by the public sector and voluntarily making data owned by natural persons and businesses available for the wider common good ("data altruism"). Building uniform European data spaces has become one of the essential components of the EU project aimed at facilitating data exchange between businesses and the private and public sectors.

The new regulations contained in Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the reuse of public sector information (OJ EU of 2019 L 172, p. 56, hereinafter: the Reuse Directive) and the Data Governance Act (COM/2020/767 final), significantly change the rules on the digital content-sharing market. The latter piece of legislation aims to establish a framework encouraging the enhanced reuse of data by increasing trust in data intermediaries, and by strengthening

data sharing mechanisms across the EU. The law will play a key role in enabling and guiding the creation of EU-wide common interoperative data spaces in strategic sectors, including those using Artificial Intelligence (AI).

The proposal lays down the rules applicable to the conditions for the reuse of protected public-sector data, including in relation to commercially confidential data, intellectual property, and data protection, as well as the obligations of data sharing providers, defined as entities providing various types of intermediary services. It also introduces the concept of data altruism and the possibility of registering an organisation as a "Data Altruism Organisation", recognised in the EU. Another proposal was to establish the European Data Innovation Board, a new formal group of experts headed by the European Commission. It is worth noting that "data" are defined in the draft as "any digital representation of acts, facts, or information, and any compilation of such acts, facts, or information, including in the form of sound, visual, or audiovisual recording". It is a broad definition which includes personal data as set out in the GDPR. Therefore, the GDPR and the Act can apply simultaneously. The explanatory memorandum also states that "measures are designed in a way that fully complies with data protection legislation and strengthens in practice the control natural persons have over the data they generate".

5 Notice and takedown in the business of online content-sharing service providers

The issue of sharing digital content online concerns primarily the relationships between online content-sharing service providers and data subjects with regard to intellectual property, as outlined in the Directive on Copyright in the Digital Single Market. It can be assumed that the foremost responsibility of rightsholders is to provide online content-sharing providers with the information they need to identify content. Providers, in turn, should ensure transparency for the implemented identification and follow-up measures. When assessing the proportionality and effectiveness of the implemented measures, technological constraints and other difficulties should be taken into account, as should the number and type of works or other protected subject-matter shared by content users. Under Article 15 of Directive 2000/31/EC, the implementation of measures by service providers should not consist of a general monitoring obligation but should be limited to ensuring the non-availability of unauthorised uses of their services of specific and duly notified copyright-protected works or other subject-matter.

Thus, it is fundamental to maintain the balance between users' rights and rightsholders' rights to content under the Charter of Fundamental Rights of the European Union (www.eur-lex.europa.eu).

Notably, the implemented measures should not require the identification of individual users sharing content online and should not involve the processing of data about individual users under the GDPR and Directive 2002/58/EC. In particular, where a given

content is subject to sharing restrictions, online content-sharing service providers should be required to provide a complaint mechanism intended for users whose data integrity has been compromised. Such a mechanism should allow users to determine why certain content is targeted by these measures and make it easier to find basic information about notable exceptions and applicable limitations.

If authors or performers issue a licence or transfer rights, they expect their work or performance to be exploited. It happens, however, that works or performances which have been licensed or transferred are not exploited at all. And when these rights have been transferred on an exclusive basis, authors and performers cannot turn to another partner to exploit their work. In such a case, and after a reasonable period of time has elapsed, authors and performers should have the right of revocation. Revocation should also be possible when the transferee or licensee has not complied with their reporting or transparency obligation, as provided for in Article 14 of the Directive on Copyright in the Digital Single Market. Revocation should only be considered after all the steps of alternative dispute resolution have been completed, particularly concerning reporting. As exploitation of works can vary depending on the sectors, specific provisions could be taken at the national level to reflect the specificities of the sectors, such as the audiovisual sector, or the works and the anticipated exploitation periods, notably providing for time limits for the right of revocation. Online content-sharing service providers perform an act of communication to the public, or an act of making content available to the public, for the purposes of this Directive, when they give the public access to copyright-protected works or other protected subject-matter uploaded by users. Hence, liability arising from dissemination will also apply to content which remains outside the control of, and is not moderated by, the provider. Therefore, an online content-sharing service provider must obtain authorisation from the rightsholders, for instance, by concluding a licensing agreement, to disseminate works or other protected subject-matter or to make them available to the public. When authorisation is obtained by way of a licensing agreement, that authorisation also covers acts carried out by users of the services falling within the scope of Article 3 of Directive 2001/29/EC, when they are not acting on a commercial basis or where their activity does not generate significant revenues. When an online content-sharing service provider performs an act of communication to the public or an act of making content available to the public, the limitation of liability established in Article 14(1) of Directive 2000/31/EC does not apply to situations involving copyrights. This rule, however, applies to providers of services beyond the scope of the Directive on Copyright in the Digital Single Market. If no authorisation is granted, online content-sharing service providers are liable for unauthorised acts of communication to the public, including making copyright-protected works and other subject-matter available to the public, unless the service providers demonstrate that they have a) made best efforts to obtain an authorisation, and b) made best efforts, following high industry standards of professional diligence, to ensure the unavailability of specific works and other subject-matter for which the rightsholders have provided the service providers with the relevant and necessary information, and c) in any event acted expeditiously, upon receiving a

sufficiently substantiated notice from the rightsholders, to hinder access to, or to remove from their websites, the notified works or other subject-matter, and made best efforts to prevent their future uploads.

Under Article 22 of the Directive on Copyright and Related Rights in the Digital Single Market, where an author or a performer has licensed or transferred their rights in a work or other protected subject-matter on an exclusive basis, the author or performer may revoke, in whole or in part, the licence or the transfer of rights where there is a lack of exploitation of that work or other protected subject-matter.

In determining whether the service provider has complied with these obligations, the following should be particularly taken into account: a) the type, audience and size of the service, and the type of works or other subject-matter uploaded by the users of the service; and b) the availability of suitable and effective means, and their cost for service providers.

For new online content-sharing service providers whose services have been available to the public in the Union for less than three years, and which have an annual turnover below EUR 10 million, calculated under Commission Recommendation 2003/361/EC (20) (OJ L 124, 20/05/2003, p. 36), the conditions under the liability regime set out in Paragraph 4 are limited to compliance with Point (a) of Paragraph 4, and to acting expeditiously, on receiving a sufficiently substantiated notice, to hinder access to the notified works or other subject-matter, or to remove those works or other subject-matter from their website.

Where the average number of monthly unique visitors of such service providers exceeds 5 million, calculated based on the previous calendar year, they shall additionally demonstrate that they have made best efforts to prevent further uploads of the notified works and other subject-matter for which the rightsholders have provided relevant and necessary information.

Due to the right to communication and freedom of expression, cooperation between online content-sharing service providers and rightsholders might not prevent works or other subject-matter uploaded by users, which do not infringe copyright and related rights, from being available. This includes such works or other subject-matter that are covered by an exception or limitation. Users in each Member State should be able to rely on any of the following existing exceptions or limitations when uploading and making available content generated by users on online content-sharing services: a) quotation, criticism, review; and b) use for caricature, parody or pastiche purposes. Online content-sharing service providers are required to provide rightsholders, at their request, with adequate information on the functioning of their practices concerning the cooperation referred to in Paragraph 4 and, where licensing agreements are concluded between service providers and rightsholders, with information on using content covered by the agreements. Furthermore, they must put in place an effective and expeditious complaint and redress mechanism available to users of their services in the event of disputes over the disabling

of access to, or the removal of, works or other subject-matter uploaded by them. Where rightsholders request to have access to their specific works or other subject-matter disabled, or to have those works or other subject-matter removed, they must duly justify the reasons for their requests. Complaints submitted under the mechanism envisaged in the first subparagraph must be processed without undue delay, and decisions to hinder access to or remove uploaded content are subject to human review. Member States also ensure that out-of-court redress mechanisms are available for settling disputes. Such mechanisms must enable disputes to be settled impartially and may not deprive users of the legal protection afforded by national law, without prejudice to the users' rights to have recourse to efficient judicial remedies. In particular, Member States must ensure that users have access to a court or another relevant judicial authority to assert the use of an exception or limitation to copyright and related rights.

Specific provisions for the revocation mechanism may be provided for in national law, taking into account the following: a) the specificities of the different sectors and the different types of works and performances and where a work or other subject-matter contains the contribution of more than one author or performer, the relative importance of the individual contributions, and the legitimate interests of all authors and performers affected by the application of the revocation mechanism by an individual author or performer. Member States may exclude works or other subject-matter from the application of the revocation mechanism if such works or other subject-matter usually contain contributions of a plurality of authors or performers. Member States may further provide that the revocation mechanism can only apply within a specific time frame, where such restriction is duly justified by the specificities of the sector or the type of work or other subject-matter concerned. Member States may provide that authors or performers can choose to terminate the contract exclusivity instead of revoking the licence or transfer of the rights.

The author or the performer must notify the person to whom the rights have been licensed or transferred and set an appropriate deadline by which the exploitation of the licensed or transferred rights is to take place. After the expiry of that deadline, the author or the performer may choose to terminate the contract exclusivity instead of revoking the licence or transfer of the rights. Member States may provide that any contractual provision derogating from the revocation mechanism is enforceable only if it is based on a collective bargaining agreement.

As noted by G. Frosio and S. Mendis, Article 17(9) of the Directive represents the legislative culmination of a global trend that inclines towards the implementation of digital content monitoring and filtering systems by intermediaries, digital content providers, meaning the transformation of the role and status of digital service providers as being liable for the content made available. Such a shift in roles in the digital market requires a fair balance between the diverse interests of users and intermediaries in disseminating and using copyright-protected content online (Frosio, Mendis, 2020:565).

Issues around the new regulations on digital content-sharing online, including the liability of digital service providers, have long been the subject of inquiry into the new field of exploitation of works. There has been a tendency in recent years to authorise digital service providers to pre-monitor all content uploaded by users. This seems to be a precondition for such content to be used appropriately. Providers could be absolved from direct and indirect copyright liability on condition that they could be shown to have implemented the appropriate content recognition and filtering technology to counter online infringements of copyrights. It involves "notice and staydown" responsibilities, where regular notifications of copyright subjects about removing illegal files would entail the obligation of proactively identifying and eliminating any instances of content purported to violate the law, and preventing the upload of such content in the future. It should be mentioned that the EC has officially confirmed, in its Communication on tackling illegal content online COM(2017) 555 final), that providers should "voluntarily" fulfil these obligations and that their scope should not be limited to copyright issues but should also include identifying and removing illegal material, such as terrorist and hate speech material. Online platforms may become aware of illegal content in several different ways, through different channels. Such channels for notifications include (i) court orders or administrative decisions; (ii) notices from competent authorities (e.g., law enforcement bodies), specialised "trusted flaggers", intellectual property rightsholders or ordinary users; and (iii) the platforms' investigations or knowledge.

Similar measures were applied under provider self-regulation schemes. An example is Google rules, under which rightsholders should upload content by sharing reference files with metadata in order to use Content ID. Google explained that even small elements of content use could be detected, regardless of whether or not any significant modifications had been made. Under YouTube's business model, rightsholders may prevent the display of copyright-protected materials on YouTube to control how their content is used, without being able to take any other measures or to make a profit from advertisements accompanying their content. New content uploaded to YouTube is checked on a fingerprint database, and then YouTube implements business rules to protect rightsholders. This video-sharing platform also has a policy in place for notifying illegal content. YouTube has several Policies on digital content sharing on its platform.

"You might not like everything you see on YouTube. If you think content is inappropriate, use the flagging feature to submit it for review by our YouTube staff. Our staff carefully reviews flagged content 24 hours a day, 7 days a week, to determine whether there's a violation of our Community Guidelines. Our products are platforms for free expression. But we don't support content that promotes or condones violence against individuals or groups based on race or ethnic origin, religion, disability, gender, age, nationality, veteran status, caste, sexual orientation, or gender identity, or content that incites hatred based on these core characteristics. It's not okay to post violent or gory content that's primarily intended to be shocking, sensational, or gratuitous. If posting graphic content in a news

or documentary context, please be mindful to provide enough information to help people understand what's going on in the video. Don't encourage others to commit specific acts of violence. YouTube is not for pornography or sexually explicit content. If this describes your video, even if it's a video of yourself, don't post it on YouTube. Also, be advised that we work closely with law enforcement, and we report child exploitation. Don't post videos that encourage others to do things that might cause them to get badly hurt, especially kids. Videos showing such harmful or dangerous acts may get age-restricted or removed depending on their severity. Everyone hates spam. Don't create misleading descriptions, tags, titles, or thumbnails to increase views. It's not okay to post large amounts of untargeted, unwanted or repetitive content, including comments and private messages. It's not ok to post abusive videos and comments on YouTube. If harassment crosses the line into a malicious attack, it can be reported and may be removed. In other cases, users may be mildly annoying or petty and should be ignored. Respect copyright. Only upload videos that you made or that you're authorised to use. This means abstaining from uploading videos you didn't make or using content in your videos that someone else owns the copyright to, such as music tracks, snippets of copyrighted programs, or videos made by other users, without necessary authorisations.

Things like predatory behaviour, stalking, threats, harassment, intimidation, invading privacy, revealing other people's personal information, and inciting others to commit violent acts or to violate the Terms of Use, are taken very seriously. Anyone caught doing these things may be permanently banned from YouTube.

If someone has posted your personal information or uploaded a video of you without your consent, you can request removal of content based on our Privacy Guidelines. Accounts that are established to impersonate another channel or individual may be removed under our impersonation policy. Learn about how we protect minors in the YouTube ecosystem. Also, be advised that we work closely with law enforcement, and we report child endangerment".

As suggested by these documents, on the one hand, the platform presents itself explicitly as an intermediary. On the other, in addition to self-regulation, the issues mentioned above are, to some extent, addressed by the Audiovisual Media Services Directive (OJ L 95, 15/4/2010, pp. 1–24), laying down the rules of liability for video sharing platform operators.

The binding rules for protecting intellectual property online are diverse and highly fragmented. For example, Google has reached an agreement with the French audiovisual industry to provide rightsholders with direct access to content removal and blocking tools on YouTube. Yet, the list of the most popular torrent sites, compiled to track the popularity of these websites, is still long. Some of them contain links to adware. As of early 2020, the most popular torrent sites were The Pirate Bay, YTS.lt (dedicated to globally popular film productions; YTS has been recently litigated against in three cases

in the US), 1337x, RARBG (the site operates from several popular domains, but only the one with the highest traffic is listed); Torrentz2, EZTV.io, LimeTorrents, Fitgirl, and Tamilskie Rockers.

As regards the assessment of copyright protection solutions for digital content, the implementation of the procedures provided for in the Directive on Copyright in the Digital Single Market would be compatible with the right of online platforms, including in particular social networking sites, to establish the framework of liability, and with the right of Internet users to fair trial, privacy, and freedom of expression under Articles 6, 8 and 10 of the 1950 European Convention on Human Rights (ECHR) (Polish Journal of Laws of 1993, No. 61, item 284), i.e., the right to a fair trial, respect for private and family life, and freedom of expression.

The key question is whether suspension and refusal-of-access procedures for digital content are compatible, in particular with the right to hold opinions, and to receive and impart information and ideas without interference by public authorities, regardless of frontiers. The lawfulness of implementing a notice and staydown regime would largely depend on whether the technology in question conforms with the three-step test performed by the ECHR in Strasbourg. In accordance with the ECHR, any interference with Articles 8 and 10 must be "in accordance with the law", serve one or more of the legitimate interests referred to in Article 8(2) and Article 10(2), and be both "necessary" and "proportionate". The exercise of these freedoms, since it carries with it duties and responsibilities, may be subject to such formalities, conditions, restrictions or penalties as are prescribed by law and are necessary in a democratic society, in the interests of national security, territorial disorder or crime, for the protection of health or morals, for the protection of the reputation or rights of others, for preventing the disclosure of information received in confidence, or for maintaining the authority and impartiality of the judiciary.

6 Issues around notice and takedown, and the right to privacy, freedom of speech, and ownership rights

In 1788, James Madison, a co-author of the US Constitution, wrote, "If men were angels, no government would be necessary. If angels were to govern men, neither external nor internal controls on government would be necessary. In framing a government which is to be administered by men over men, the great difficulty lies in this: you must first enable the government to control the governed, and in the next place oblige it to control itself". And this thought remains as relevant as ever concerning future social media regulations.

It should be noted that the adoption of digital content recognition and filtering technology, by its very nature, raises serious concerns about not only the right to privacy but also the right to freedom of expression. First, content notice and staying down might lead to issues with the right to privacy. This concerns, for instance, social media users who illegally

upload copyrighted content through a social networking platform and disable access to such content (www.internetsociety.org/about-internet-society/annual-review/2010). This stems from the fact that content recognition and filtering generally rely on fingerprinting technology, watermarks, real-time monitoring, and identifying illegal user content before blocking access. In particular, unlike blocking measures, DPI (deep packet inspection) systems are a type of data processing that inspects in detail the data being sent over a computer network and can take actions such as alerting, blocking, re-routing, or logging it accordingly. Deep packet inspection is often used to baseline application behaviour, analyse network usage, troubleshoot network performance, ensure that data are in the correct format, check for malicious code, eavesdropping, and Internet censorship (Duncan Geere, https://www.wired.co.uk/article/how-deep-packet-inspection-works), and investigate network packets instead of focusing on the source, such as a URL blacklist. The DPI technology can reveal information which makes it easy to establish user identity, location, interests, activities, etc. The general obligation to retain data is, however, incompatible with the personal data protection system unless it fulfils specific conditions. See an opinion of Advocate General Henrik SaugmandsgaardØe, delivered on 19 July 2016: "Mr Schrems, an Austrian national residing in Austria, is a user of the social network Facebook. All users of that social network residing in the territory of the European Union are required, when signing up, to enter into a contract with Facebook Ireland, a subsidiary of Facebook Inc., which is established in the United States. Those users' personal data are transferred, in whole or in part, to servers belonging to Facebook Inc., situated in the territory of the United States, where they are processed. On 25 June 2013, Mr Schrems filed a complaint with the DPC whereby he requested her, in essence, to prohibit Facebook Ireland from transferring the personal data relating to him to the United States. He claimed that the law and practices in force in the United States did not ensure adequate protection of the personal data retained in its territory against intrusions resulting from the surveillance activities practised by the public authorities. Mr Schrems referred in that regard to the revelations made by Mr Edward Snowden concerning the activities of the United States intelligence services, in particular those of the National Security Agency (NSA)" (Case C-311/18).

In that regard, the Court held, in the judgment in Schrems (see: C-203/15 and C-698/15 Tele2 Sverige AB v. Post-ochtelestyrelsen (C-203/15) and Secretary of State for the Home Department v. Tom Watson, Peter Brice, Geoffrey Lewis (C-698/15), with the participation of Open Rights Group, Privacy International, Law Society of England and Wales) that the legislation which does not provide for any possibility for an individual to pursue legal remedies in order to have access to personal data relating to him (Article 15 of the GDPR, entitled "Right of access by the data subject," stipulates in Paragraph 1 that "The data subject shall have the right to obtain from the controller confirmation as to whether or not personal data concerning him or her are being processed, and, where that is the case, access to the personal data [...]". The "access principle" provided for in Annex II (II) (a) of the Privacy Shield has the same underlying purpose) or to obtain the rectification or erasure of such data, does not respect the essence of the fundamental right

enshrined in Article 47 of the Charter. It must be emphasised that that right of access entails the possibility for a person to obtain from the public authorities, subject to the derogations that are strictly necessary to pursue a legitimate interest, a confirmation of whether they are or are not processing data of a personal nature relating to him or her. And it does not matter here whether the person concerned is unaware of whether the public authorities have retained personal data relating to him or her, following, inter alia, an automated filtering process of electronic communications flows.

What's more, notably, J. Urban, J. Karaganis, and B. Schofield claim that notice and takedown also raise concerns as to freedom of expression, since user choices are made by algorithm matching based on big data, and not only in a context understandable exclusively to humans (Urban, Karaganis, Schofield, 2016:8). What is interesting is that each week Google receives millions of take-down requests, sent mainly by or on behalf of major entertainment corporations (Urban, Karaganis, Schofield, 2016:11). Thus, the use of this technology might easily lead to mistakes, especially to blocking legal content (falsely positive), or sharing illegal materials (falsely negative). It is emphasised that the right to freedom of expression might be infringed when exemptions apply which are not protected under copyright laws or laws on personal data protection, e.g., when content belongs to a public domain or is misdetected and removed, leading to "false results", as mentioned above. What is also important is that some systems are highly efficient in recognising content but targeting illegal remixes, DJ sets, and mashups can still be very difficult in the case of copyright-protected content and other content which features hate speech that is not explicit but contextual.

The ECHR also noted these circumstances in the case Delfi v. Estonia, where the Court held that "Delfi's news portal had a disclaimer stating that the writers of the comments – and not the applicant company – were accountable for them and that the posting of comments that were contrary to good practice or contained threats, insults, obscene expressions or vulgarities, or incited hostility, violence or illegal activities, was prohibited. Furthermore, the portal had an automatic system of deletion of comments based on stems of certain vulgar words, and it had a notice-and-take-down system in place, whereby anyone could notify it of an inappropriate comment by simply clicking on a button designated for that purpose to bring it to the attention of the portal administrators. On some occasions, the administrators removed inappropriate comments on their own initiative. Thus, the applicant company could not be said to have wholly neglected its duty to avoid causing harm to third parties. Nevertheless, and more importantly, the automatic word-based filter used by the applicant company failed to filter out odious hate speech and speech inciting violence posted by readers, and thus limited its ability to expeditiously remove the offending comments. Notably, the majority of the words and expressions in question did not include sophisticated metaphors or contain hidden meanings or subtle threats. They were manifest expressions of hatred and blatant threats to the physical integrity of the injured party. Thus, even if the automatic word-based filter may have been useful in some instances, the facts of the present case

demonstrate that it was insufficient for detecting comments whose content did not constitute protected speech under Article 10 of the Convention". Google similarly claimed that imposing the obligation to notify and block service providers' content was illegitimate. Some mechanisms, also concerning discretion in their use, can be used in one context, such as Content ID or YouTube, but not in another, e.g., social networking platforms (Facebook, Twitter, Snapchat, Instagram, etc.). Google explains that it is still easier than ever for authors to connect with their audiences, build fanbases, and share their content online using these networking platforms (EC 2016b, 164–165).

The DSA states, "Union citizens and others are exposed to ever-increasing risks and harms online. According to the EU legislators, the Digital Services Act introduces important safeguards to allow citizens to freely express themselves while enhancing user agency in the online environment, as well as the exercise of other fundamental rights such as the right to an effective remedy, non-discrimination, rights of the child as well as the protection of personal data and privacy online. The proposed Regulation will mitigate risks of erroneous or unjustified blocking of speech, address the chilling effects on speech, and stimulate the freedom to receive information and hold opinions. The proposal will only require the removal of illegal content and will impose mandatory safeguards when users' information is removed, including the provision of explanatory information to the user, complaint mechanisms supported by the service providers, as well as external out-of-court dispute resolution mechanisms. Furthermore, it will ensure EU citizens are also protected when using services provided by providers not established in the Union but active on the internal market since those providers are covered too. Obviously, this applies, above all, to major social networking platforms. Therefore, an important objective of the DSA was to introduce uniform notice and action (notice and takedown) mechanisms across the EU" (COM(2020) 825 final).

Hence, it is imperative to provide clear and user-friendly mechanisms for users to notify or flag illegal content to intermediaries. Online intermediaries have been obliged to verify the notified content, and to respond to the notifying party and content provider within a reasonable time, explaining why the notified content had to be blocked or removed or why it remained online. It is critically important to put procedures in place for intermediaries to appeal against moderation decisions. What is important is that the steps taken by online content-sharing service providers, cooperating with rightsholders, should be without prejudice to the application of exceptions or limitations to copyright, particularly those which guarantee the freedom of user expression. Users should be allowed to upload and make available the content generated by users for the specific purposes of quotation, criticism, review, caricature, parody or pastiche. That is particularly important for striking a balance between the fundamental rights laid down in the Charter of Fundamental Rights of the European Union, in particular the freedom of expression and the freedom of the arts, and the right to property, including intellectual property. In judgments in the cases Sabam v. Netlog and Sabam v. Scarlet, the CJEU

refused to impose on those service providers the obligation to automatically monitor the content disseminated by their users under Articles 8, 11, and 16 of the Charter.

Digital content moderation alone is problematic from a regulatory standpoint. Indeed, as noted by N. Elkin-Koren and M. Perel, it blurs the distinction between private interests and public responsibilities, delegates the power to make social choices about content legitimacy to obscure algorithms, and circumvents the constitutional safeguard of the separation of powers (Elkin-Koren, Perel, 2020:671). A prime example of this was seen when Facebook and Twitter blocked the accounts of the then-incumbent US President Donald Trump or when they took down historical content about children's concentration camps in Łódź for political reasons. Another instance involved the blocking of a short animation produced by the Polish Institute of National Remembrance (IPN) concerning Poland's modern history ("Niezwyciężeni" – "The Unconquered"), which premiered on 15 September 2010. Commissioned by IPN, the video was made in two languages, Polish and English. YouTube recently blocked the latter due to "a copyright claim". "The Unconquered" is about Poland's history from World War II to the fall of the Iron Curtain. Its Polish version on YouTube was viewed almost 1.2 million times (and received 68,000 likes and 1,000 dislikes). The animation video became popular immediately after its première, not only among Polish users. On IPN's profile alone, it had nearly 35,000 views. It took one click to remove a video having more than 2 million views on Facebook from IPN's YouTube account. M. Poślad, Head of CEE & Transatlantic Public Policy at Google, wrote that the platform was legally obliged to take down the material.

Marta Poślad
@MartaPoslad
YouTube is required by law to block notified videos. Awaiting the IPN's response to the copyright claim https://goo.gl/jQEEYa

niezależna

Niezalezna.PL
@niezaleznapl
YouTube blocks video on Poles' heroism. The Internet is on fire http://fb.me/BYmhTAs2
4:26 PM · 23 Sept. 2017

It should be emphasised that, in the case Sabam v. Netlog (OJ C 98, 31.3.2012, pp. 6–7), the CJEU examined whether or not requiring a social networking platform which shared third-party content to use notice and takedown measures to enforce copyrights online represented an infringement of fundamental rights. The CJEU concluded that the identification, systematic analysis, and processing of personal information connected with the profiles of Netlog users could represent an infringement of their right to privacy in the context of Article 8 of the Charter of Fundamental Rights. This judgment also reiterates CJEU's previous decision in the case Sabam v. Scarlett (Case C-70/10), and later it was reiterated in CJEU's ruling in the Mc Fadden case (Case C-484/14). Moreover, the objectives which, under that provision, such measures must pursue, such as safeguarding national security, defence, and public security, and the prevention, investigation, detection and prosecution of criminal offences or unauthorised use of the electronic communications system, overlap substantially with the objectives which legitimise restrictions on the rights and freedoms set out in Article 31(3) of the Constitution of the Republic of Poland. On the one hand, the retention of communications data enables the government to control the governed by providing the competent authorities with a means of investigation which might prove useful in fighting serious crime, particularly in combating terrorism. In substance, the retention of communications data gives the authorities a certain ability to examine the past by accessing data relating to user communications. However, on the other hand, it is imperative to oblige the government to control itself with respect to both the retention of data and access to the data retained, given the grave risks engendered by the existing databases encompassing all communications made within the national territory. Indeed, these enormous databases give anyone with access the power to instantly catalogue every member of the population in question. These risks must be scrupulously addressed, in particular through examining the strict necessity and proportionality of the general obligation for digital content providers to block and remove content, and to provide data to public authorities.

Accordingly, it is necessary to strike a fair balance between the obligation of Member States to ensure the protection of individuals on their territory, the observance of the fundamental rights to private life and personal data protection, and the protection of intellectual property. According to ECHR's case law, for any interference with the right to privacy or freedom of expression to be "lawful" under Articles 8 and 10 of the Convention, the following three conditions must be satisfied. First, it must be based on domestic legislation; second, such legislation should be accessible; and third, the legislation should follow the Strasbourg Court's rules of predictability and legality. The adoption of notice and action systems could be incompatible with the Court's requirements regarding accessibility, predictability, and legality, thereby violating the first part of its non-cumulative test under Articles 8(2) and 10(2) of the Convention. As regards the first rule, following ECHR's case law, the quality of a legal requirement, under Articles 8 and 10 of the Convention, requires that a law be published and, by extension, that fair access to it be provided to those affected by the law. As noted above,

the assessment of the consequences concluded that, in the case at hand concerning copyright protection, rightsholders must supply service providers with the information necessary for content identification, and these services must furnish rightsholders with the "appropriate information" about systems.

What is concerning, however, is that users are refused access to the technical details of these evidence collection techniques, as a result of which they may not rely on the judicial review to question their use. In the case Sabam v. Scarlet, the Court observed that notice and takedown involved filtering all electronic communication passing through the ISP to identify individuals engaged in copyright infringement, as well as blocking all incoming and outgoing communication involving such an infringement. The reference for a preliminary ruling concerned the dispute between the company Scarlet and Sabam, a Belgian society of authors, composers, and publishers. It concerned Scarlet's refusal to install a system for filtering electronic communications which use peer-to-peer software to prevent file sharing, which infringes copyright. The national court asked the question of whether EU regulations permit national courts to be authorised to issue an injunction against intermediaries, for all its customers, *in abstracto*, and as a preventive measure, exclusively at the cost of that intermediary and for an unlimited period, a system for filtering all electronic communications, both incoming and outgoing, passing via its services, to identify on its network the movement of electronic files containing works in respect of which the applicant claims to hold rights, and subsequently to block the transfer of such files. The Court held that such an injunction would result in a serious infringement of the freedom of the ISP concerned to conduct its business. Furthermore, it would not respect the requirement that a fair balance be struck between protecting intellectual property rights and protecting the freedom to conduct business enjoyed by operators such as ISPs. Lastly, Scarlet claimed that installing a filtering system would be in breach of the provisions of the European Union law on the protection of personal data and communications secrecy, since such filtering involves the processing of IP addresses, which are personal data. In that context, the referring court regarded that, before ascertaining whether a mechanism for filtering and blocking peer-to-peer files existed and could be effective, it had to be satisfied that the obligations liable to be imposed on Scarlet were in accordance with the European Union law. Accordingly, such an injunction would result in a grave infringement of the freedom of the ISP concerned to conduct its business since it would require that ISP to install a complicated, costly and permanent computer system at its own expense, which would also be contrary to the conditions laid down in Article 3(1) of Directive 2004/48, requiring that measures to ensure the respect of intellectual property rights should not be unnecessarily complicated or costly.

As mentioned before, the assessment of consequences concluded that rightsholders claimed that the functioning of such technologies remains largely "unclear" to them (EC 2016a, 141). As regards the predictability rule, in accordance with ECHR case law, there must be a sufficient degree of predictability in law as to the scope of the applicable measures, as guaranteed by Articles 8 and 10 of the Convention. Worryingly, this

suggests that the level of control required to implement this technology constitutes an intrusive analysis of both personal and sensitive data. Hence, since notice, staydown, and takedown depend on the monitoring equipment, the level of investigation required to monitor users must be clearly defined. Another pertinent issue is whether installing content recognition and filtering would pass the third part of the ECHR's three-step test. According to ECHR case law, under Articles 8(2) and 10(2) of the Convention, supervisory and technical measures are "necessary" in a democratic society if they address "an urgent social need" and are proportionate means of achieving a legitimate aim. Furthermore, the ECHR noted that the state's explanation of such measures must be "adequate and relevant" although state authorities have a certain margin of discretion.

In reality, this goes along the lines of the judgment in the Delfi v. Estonia case, in which the ECHR held that, if accompanied by effective procedures allowing for rapid response, the notice and takedown would represent an appropriate tool for balancing the rights and interests of all those involved. In its judgment on the case Sabam v. Netlog, the CJEU explained that the notice and staydown solution constituted a breach of EU law since it required social networking platforms to implement filtering technology for all communications. Moreover, technology cannot handle complex decisions such as determining whether a certain use is lawful, identifying copyright ownership, and avoiding mistakes, duplicates, or overblocking (Urban, Karaganis, Schofield, 2016:35). In its judgment in the case Sabam v. Netlog, the CJEU noted that Article 15(1) of the Electronic Commerce Directive prohibits national judges from imposing general monitoring obligations on social networking platforms. According to the CJEU, because these platforms were required to implement a complex, expensive, and permanent system, their freedom to conduct business was affected significantly. In particular, it found that such technology violated Article 3(1) of Directive 2004/48/EC (OJ L 157, 30/04/2004, pp. 45–86). Filtering also involved the detection, automated analysis, and processing of personal data, likely blocking legal communications. Relying on the case Promusicae v.Telefonica (Case C-275/06), the CJEU concluded that notice and staydown did not result in a fair balance between the rightsholders right to intellectual property, on the one hand, and the freedom to conduct business by other social networking platforms, as well as the users' rights to personal data protection and to receive and impart information, on the other.

Regarding the obligations imposed on intermediaries, it is vital to recall the ECHR's decision in the case (40397/12) Neij &SundeKolmisoppi v. Sweden. During 2005 and 2006, Fredrik Neij and Peter SundeKolmisoppi were involved in running one of the world's largest file-sharing (music, movies, computer games) services on the Internet – *The Pirate Bay* (TPB). In 2008, they and others were charged with complicity in committing crimes in violation of the Copyright Act. As a result, several companies in the entertainment business brought private claims against them. In April 2009, the District Court in Stockholm sentenced them to one year's imprisonment and held them jointly liable for damages of approximately EUR 3.3 million, together with the other defendants.

In November 2010, the Court of Appeal in Svea reduced their prison sentences but increased their liability for damages to approx. EUR 5 million. In their application to the Court, both defendants argued that they were not liable for how other individuals used the TPB website, whose original purpose was to facilitate online data sharing. They claimed that crimes were being perpetrated only by those users who had exchanged illegal information about copyrighted material.

Accordingly, in reliance on Article 10 of the Convention, they argued that their conviction for complicity in committing crimes in violation of the Copyright Act represented an infringement of their right to freedom of expression (Ombelet, Kuczerawy, Valcke, 2016: 4). The Court found that Article 10 of the Convention guaranteed the right to impart information and the public's right to receive it. In light of its accessibility and capacity to store and communicate vast amounts of information, the Internet plays a significant role in enhancing public access to news and in facilitating the sharing and dissemination of information. Moreover, it applies not only to the content of the information but also to the means of transmission or reception since any restriction imposed on the means necessarily interferes with the right to receive and impart information. Further, Article 10 guarantees freedom of expression to "everyone". No distinction is made in it according to whether or not the aim pursued is profit-making. The Court found that the Swedish authorities were obligated to protect the plaintiffs' property rights under the Copyright Act and the Convention, and that there were weighty reasons for restricting the applicant's freedom of expression.

Moreover, the Swedish courts advanced relevant and sufficient reasons to consider that the applicant's activities within the commercially run TPB amounted to criminal conduct requiring the appropriate punishment. "In this respect, the Court reiterates that the applicants were only convicted for copyright-protected materials. In reaching this conclusion, the Court has regard to the fact that the domestic courts found that the applicants had not taken any action to remove the torrent files in question despite having been urged to do so. Instead, they had been indifferent to the fact that copyright-protected works had been the subject of file-sharing activities via TPB". Consequently, the Court also found that, due to the nature of the information contained in the shared material and the weighty reasons for the interference with the applicant's freedom of expression, this interference was "necessary in a democratic society", within the meaning of Article 10(2) of the Convention.

On the one hand, A. Lucas-Schotter claims that Article 13 of the Directive on Copyright in the Digital Single Market is a well-balanced text which, despite attracting sharp criticism, is fully compliant with Community laws and does not violate the Charter of Fundamental Rights of the European Union and the Electronic Commerce Directive (Lucas-Schoetter, 2017:21). On the other hand, there has been a growing number of disputes over content recognition and filtering systems. Ch. Angelopoulos and S. Smet opined that, using the example of copyright, the risk exists that no resolution would be

possible. "When two industries with conflicting interests are asked to self-regulate, it only entrenches the differences in their business models, and that is why "cooperation" between Internet service providers and the entertainment industry struggles to work without a court ruling" (Angelopoulos, Smet, 2016:301; Horten, 2016:142).

It is worth concluding that regulations related to the notice and action obligation, although supported by the ECHR's ruling, are criticised for completely disregarding the role of service providers and digital media in society. In the case Delfi v. Estonia, judges Sajó and Tsotsoria (Case 64569/09) observed that, in cases where an individual victim exists, they may be prevented from notifying an Internet service provider of the alleged violation of their rights. The Court attaches weight to the consideration that the ability of a potential victim of hate speech to monitor the Internet continuously is more limited than the capability of a large commercial Internet news portal to prevent or rapidly remove such comments. Therefore, a large news portal's obligation to take effective measures to limit the dissemination of hate speech and speech inciting violence – the issue in the present case – can, by no means, be equated to "private censorship". While acknowledging the "important role" played by the Internet "in enhancing public access to news, and facilitating the dissemination of information in general, it is also mindful of the risk of harm posed by content and communications on the Internet".

Another issue pertains to user anonymity. Internet users' interest in not revealing their identity seems critical. Anonymity has long been a means of avoiding reprisals or unwanted attention. As such, it can promote the free flow of ideas and information. At the same time, the ease, scope, and speed of the dissemination of information on the Internet, and the persistence of the information once disclosed, may considerably aggravate the effects of unlawful speech on the Internet compared to traditional media. Different degrees of anonymity are possible on the Internet. An Internet user may be anonymous to the broader public while being identifiable by a service provider through an account or contact data, which may be either unverified or subject to some verification – ranging from limited verification (for example, through activation of an account via an e-mail address or a social network account) to secure authentication, be it by the use of national electronic identity cards or online banking authentication data allowing somewhat more secure identification of the user. A service provider may also allow an extensive degree of anonymity for its users, in which case the users are not required to identify themselves at all, and they may only be traceable – to a limited extent – through the information retained by Internet access providers. The release of such information would usually require an injunction by the investigative or judicial authorities and would be subject to restrictive conditions. It may nevertheless be necessary in some cases to identify and prosecute perpetrators. Another aspect involves transferring personal data when illegal content has to be blocked. Under the EU-US Privacy Shield, the United States ensures a sufficient degree of protection for data the EU provides to the United States. The EU-US Privacy Shield is constituted by the principles issued by the US Department of Commerce on 7 July 2016, as set out in Annex II, and official declarations

and commitments contained in the documents presented in Annexes I, III-VII. For the purpose of Paragraph 1, personal data are transferred under the EU-US Privacy Shield, where they are transferred from the Union to organisations in the United States that are included in the "Privacy Shield List", maintained and made publicly available by the US Department of Commerce, under Sections I and III of the Principles set out in Annex II. Annex III A to this decision, entitled "EU-U.S. Privacy Shield Ombudsperson mechanism regarding signals intelligence", attached to the letter of the then Secretary of State John Kerry, dated 7 July 2016, contains a Memorandum laying down a new mediation procedure conducted before the Senior Coordinator for International Information Technology Diplomacy (Senior Coordinator), as appointed by the Secretary of State. Following the Memorandum, the procedure has been implemented "to facilitate the processing of requests relating to national security access to data transmitted from the EU to the United States pursuant to the Privacy Shield, standard contractual clauses (SCCs), binding corporate rules (BCRs), 'Derogations', (2) or 'Possible Future Derogations', (3) through established avenues under applicable United States laws and policy, and the response to those requests".

It is concerning that each monitoring system which has to be implemented by social networking platforms to comply with notice and takedown obligations might be used in the future to process users' analytical data for targeted display-advertising strategies. The problem is further exacerbated by the fact that DPI technology also makes it possible to modify content. Thus, the fundamental question arises as to whether it would be relatively easy, from the technical standpoint, to apply content monitoring, recognition, and filtering technology once the underlying infrastructure has been broadly implemented in corporations or public authorities, to block access to other information, thereby silently encouraging Internet censorship (Internet Society 2010:80, www.wipo.int).

According to the ECHR, the implementation of notice and takedown systems would be in line with social networking platforms and users' right to a fair trial, and the right to privacy and freedom of expression, provided that they are informed about the technical details of such systems, as well as about the *ratione personae*, *rationemateriae* and *rationetemporis* scopes of the supervisory and technical measures, which should be set out explicitly by law. Public authorities should be involved in checking and authorising systems, followed by regular audits. Also, it should be emphasised that notice and takedown are insufficient and limited to situations in which these systems can be considered essential to achieve a legitimate aim following the necessity and proportionality principles. As a rule, it is important to implement mechanisms to prevent the overblocking of digital content. Platforms should provide their users with simple mechanisms to question decisions on removing digital content they have uploaded. However, where no agreement can be reached in this manner, cases should be referred to court.

7 The right to data portability and monetisation

The underlying element analysed in this study was the digital content trading system in the economy created by large digital platforms. The analysis covered specific content categories, i.e., user-provided, user-generated, or service-provider-generated content based on user data. Considering that such content often includes personal data and copyrighted content, a multi-level and cross-sectoral approach was taken to categorise and classify its definitions, and to establish legal protection issues. The term *user-provided private content* is a general representation of a dataset which includes non-copyrighted content. Making money on private content is a reality in many business models, and it poses numerous legal problems related to privacy, consumer law, and the harmonised approach to intellectual property law and e-commerce. The first issue was to determine which data could be included in the category of user-provided content and which could not. Combining the wording of the right to data portability from the GDPR with the wording of the Directive on the Supply of Digital Content and the Directive on Copyright and Related Rights in the Digital Single Market, it can be assumed that the "supplied data" include both data provided actively and passively (i.e., generated by cookies) while "data provided actively" can be both directly provided (i.e., sent by the user) and indirectly provided (i.e., by accepting the service provider's access to certain specified information). The concept of analysing the existing obligations concerning content portability and the possible measures to facilitate content transfer do not imply attaining the objective of improving the interoperability of services by imposing additional regulatory obligations. Besides, the transfer of one's own content (e.g., different types of files, messages stored in intermediaries' resources, etc.) and data (in the sense of the GDPR) from one provider of certain services to another and cross-border transfer are two different things, within the meaning of the "portability" regulation. In the latter case, access to content and the possibility to exercise, to some extent, the "portability" of that access are closely connected with copyright based on the territoriality criterion. This is also a critical element in proprietary rights. Recital 70 of that Directive states that the consumer could be discouraged from exercising remedies lacking conformity of digital content or a digital service if that consumer is deprived of access to content other than personal data, which he/she has provided or created through the use of the digital content or digital service. To ensure that the consumer benefits from effective protection regarding the right to terminate the contract, the trader should, therefore, at the request of the consumer, make such content available to the consumer following the termination of the contract.

The right to data portability applies to all "supplied" data. To define the objective parameters at the boundary between "active" and "passive" data, a "test" based on the following three variables has been proposed: the data subject's activity in providing the data, the data subject's awareness of providing the data, and the data controller's activity as regards collecting the data.

The sharing of actively provided data should be considered a legitimate form of barter payment in exchange for digital content. The legal protection of natural persons as regards managing user-provided private content covers the right to "transfer" such data from one data controller to another, which arises from the GDPR. However, it does not cover data other than personal data. Users should have the right to "license" such data if they are rightsholders (e.g., authors). As shown earlier in this paper, the "licensing" of user-provided content has already become a reality in the most popular social networks. Nonetheless, the statutory terms of service provision, covering the scope of "licensing," are too broad, and licences are granted to service providers in exchange for access to a particular community. The solution adopted under Article 7(4) of the GDPR, i.e., the right of data subjects to grant consent, is crucial in terms of regulating the issue of private content as a form of payment other than money. This is a user-centred system based on the users' control and awareness of managing "private content". Administration should be based on two separate legal tools: licences to use user-generated content and the right to withdraw and transfer such content from one platform to another, i.e., the full enforceability of the right to data portability (possibly in combination with the right to erase data). This awareness should be based on transparent information obligations regarding the commercial purposes of this data processing, in order to respect the principle of freedom of consent and the principle of purpose limitation (Malgieri, Custers, 2017:2). Recital 24 of the Directive concerning the supply of digital content reads as follows: "Digital content or digital services are often supplied also where the consumer does not pay a price but provides personal data to the trader". Such business models are used in different forms in a substantial sector of the market. While fully recognising that protecting personal data is a fundamental right, and, therefore personal data cannot be considered a commodity, this Directive should ensure that consumers are, in the context of such business models, entitled to contractual remedies. This Directive should, therefore, apply to contracts in which the trader supplies, or undertakes to supply, digital content or digital service to the consumer, and the consumer provides, or undertakes to provide, personal data. The personal data could be provided to the trader either at the time the contract is concluded or at a later date, e.g., when the consumer gives consent for the trader to use any personal data that the consumer might upload or create with the use of the digital content or digital service. EU law on the protection of personal data provides for an exhaustive list of legal grounds for the lawful processing of personal data. This Directive should apply to any contract in which the consumer provides, or undertakes to provide, personal data to the trader. For example, this Directive should apply when the consumer opens a social media account and provides a name and email address, which are used for purposes other than solely supplying the digital content or digital service, or complying with the legal requirements. It should equally apply when the consumer gives consent for any material which constitutes personal data, such as photographs or posts which the consumer uploads, to be processed by the trader for marketing purposes. Member States should, however, remain free to determine whether the requirements for the formation, existence, and validity of a contract under national law have been fulfilled. When digital content and digital services are not supplied in exchange for the payment of

a set price, the Directive should not apply to situations in which the trader collects personal data exclusively to supply digital content or digital service, or for the sole purpose of meeting the legal requirements. Such situations can include, for instance, cases in which the registration of the consumer is required by the applicable laws for security and identification purposes. The Directive should not apply to situations where the trader only collects metadata, such as information concerning the consumer's device or browsing history, except when this situation is considered a contract under national law. It should also not apply to situations in which the consumer, without having concluded a contract with the trader, is exposed to advertisements exclusively to gain access to digital content or a digital service (Recital 25 of the Directive).

If the consumer provides the entrepreneur with personal data, the entrepreneur should meet the obligations under Regulation (EU) 2016/679. These obligations are equally applicable if the consumer pays the fee and provides personal data. On termination of the contract, the entrepreneur should refrain from any further use of content other than the personal data provided or created by the consumer when using the digital content or digital service provided by the entrepreneur. Such content may include digital images, audio files, video files or content created using mobile devices. However, the trader should have the right to continue to use the content provided or created by the consumer if that content is not useful outside the context of the digital content or service supplied by that trader, if it relates solely to the activity of the consumer, if it has been combined with other data by the trader and cannot be separated from it, or such separation requires disproportionate effort, or if it has been generated jointly by the consumer and other persons, and can still be used by other consumers.

Yet, there is a sphere of content in which the data subject does not knowingly provide personal data but actively selects content (e.g., images contained in a specific folder) and shares those pieces to generate new data on an individual. Thus, determining the owner of such content becomes problematic. The question arises as to whether it is user data based on newly-produced content, or whether it becomes the property of the entity which allowed such content to be generated and bore the costs involved in this process. The conditions in which digital content is used often create uncertainty as to whether such content carries "shared" or "observed" data and whether the use of such data extends far beyond the activities involved in their generation.

One of the key issues related to digital content protection is identifying the need for establishing new ownership rights regarding raw and non-personal data as a common good. However, should such data be considered copyrighted digital content, it would be necessary to provide access to it to entities which are primarily public-interest oriented. Accordingly, permissible public use should include, for instance, the use of data in advanced research and content derived from the automated analysis of large datasets.

Consequently, these guidelines should be drafted with due regard for the general-interest objectives to be achieved through the measures taken by video-sharing platform providers and the right to freedom of expression. However, it seems that such a regulation should be left to Member States, which must consider not only the three-step test but also their national perception of the general-interest objectives (judgments passed in the following cases: C-120/78 Cassis de Dijon, C-33/74 Van Binsbergen, C-205/04 Gouda, C-76/90 Säger, C-384/93 Alpine Investments). For the mandatory requirements doctrine to be applied, a three-step test must be passed to demonstrate that (1) there is an overriding general (public) interest; (2) the measures implemented to pursue this interest are appropriate and adequate; and (3) the measures applied to implement that interest are proportionate. The protection of digital content created by Internet users can be achieved by making these platforms rely on the principles of interoperability , transparency, and openness. This concerns handling the digital content of network users.

8 Concluding remarks

8.1 Regulation of digital content

The problem of programming content regulation in the context of changes related to digitisation processes, the wide range of issues related to regulation in this sphere (e.g., digitisation of archival resources and the digital archiving of the programming portfolio of contemporary audiovisual media, the protection of children and young people on the Internet, the protection of privacy and the security of identity on the net, as well as the protection of intellectual property, and the combating of "network piracy"), and the reuse of public sector information – all these shape the digital media market. Its regulation will be the first context for new solutions related to regulating digital content on the net, especially concerning responsibility for digital content. This need for such regulation stems from the evolving digital processes and the inadequacy of current provisions in meeting the needs arising from the so-called digital revolution. What now appears indispensable is a general approach which will also define, in a systemic manner, the scope of protection related to digital content processing in an ICT network. This issue not only relates to services provided electronically and to the solutions proposed in the regulation on digital services but also concerns broadly understood inter-sectoral cooperation.

8.2 Level of regulations

It is worth pointing out that the objectives of the new regulations include improving the detection of illegal content on Internet platforms or creating a fast track to enable state authorities to contact the platforms and permanent contact points for service providers and state authorities (content policy). However, the convergence context makes it difficult to determine whether a given sphere of activity falls within the scope of arrangements only for digital services or whether it already goes beyond that and applies to the

regulation of infrastructure, i.e., those aspects of digital processes which are currently regulated at the EU level. Content regulation (digital content regulation, including the digitisation of archives and the legality of sharing) should be a matter of national solutions, considering the specificity of a given country and, in particular, national culture specificity. Therefore, it is of utmost importance to identify the spheres in which the national policy should apply, as this will consequently translate into governance and management in the digital field, i.e., the entire system of administration of a given Member State, to which a given field is subordinated.

8.3 Open character of digital resources

The purpose of digitisation is not only to protect collections from destruction or loss but also to share them. For this reason, the digital content accumulated by the Polish archives, libraries, and museums must not only ensure safe storage conditions but also be as widely available for users as possible, free of charge, and in a form which allows their reuse for non-commercial purposes. Pursuing the use of new technologies in preparing, securing, and sharing collections is the natural course of action for every institution taking care of its collections. This leads to the dispersion of digitisation initiatives, characteristic of all European countries, and results in the dispersion of digital content, making it difficult for users to access the resources they seek. Digital collections are created and stored by many institutions throughout this country, often as thematic virtual exhibitions, occasional publications, or resources only available locally on the computer terminals of the home institution. Using popular online search engines to search such dispersed resources not only fails to provide users with the complete picture of the digitised Polish cultural heritage but is also time-consuming and inconvenient. The sharing of digital collections online is regulated by the Act on Copyright and Related Rights of 4 February 1994, which stipulates that works for which the author's economic rights have expired, i.e., 70 years after the author's death, and for co-authors, after the death of the last surviving author, may be publicly disseminated without limitation (Article 36(1)). For digitised works, this restriction relates to both the authors and the translators or illustrators of the work. Article 28 of the Act on Copyright and Related Rights allows the permitted use for free sharing by libraries, archives, and schools of reproductions of works remaining under copyright protection, but only on the premises of such facilities, whilst the sharing by libraries, archives, and schools of digital reproductions on the Internet is no longer a permitted use within the meaning of that Act. This largely restricts the possibilities of sharing digital collections. Digital reproductions of copyright-protected documents are thus made available only at library computer stations. Therefore, the simple role of a digital library, which is online access to electronic resources not limited by time and place, is not fulfilled. In order to open the desired item (or at least verify it, if this indeed proves desirable), the reader must visit the library, even though the remote sharing of a digital reproduction would not be a technical problem. It is also essential to identify laws which could hinder the online sharing and reuse of publicly-owned cultural material. Unfortunately, Poland still lacks solutions to this problem. It should be emphasised that

exceptions to and limitations on copyright, including the principles of *fair use* and *fair dealing*, ensure an effective balance between the protection of authors in the scope of their creative activity, resulting from copyright or related rights, and the public interest. Such mechanisms guarantee certain privileges to users. This, in turn, opens up free space for action in the current copyright system. Given the rapid changes occurring in the field of technology and social behaviour, it is imperative to ensure the possibility of action using legally protected resources. Any restrictions on copyright, *fair use,* and *fair dealing* should be flexible and constantly adjusted to the needs and goals of the public interest.

It should also be noted that crucial questions have so far arisen about the limits of subjecting content to infrastructure regulation when the issue of market regulation is dominant. Yet, it seems there has been a new trend towards the reverse situation, whereby the regulation of infrastructure is subjected to digital content regulation. This also applies to the public sector and the information generated, stored, and processed there. Property rights can have an impact on restrictions pertaining to this subject. Such information, subject to the intellectual property rights of third parties, may not be reused. Therefore, the entity must refuse to reuse public-sector information if the intellectual property right does not belong to that entity or if it only has the right to use the work. Correspondingly, this applies to resources covered by legally protected secrets and to resources owned by network users, though they are not necessarily the subject of copyright or related rights.

One of the underlying problems concerning the regulation was the issue of determining whether public resources – public-sector information – are only those financed from public funds or also include resources which are owned by social organisations or natural persons but which have been made available to the public as part of the activities of public institutions. An important issue concerned determining whether the regulation should cover the sharing of national resources. This includes non-public collections, which are in the possession and at the disposal of public-sector institutions. The most crucial postulates of the groups involved regarding the provisions on reuse were the necessity to precisely define the scope of the public domain and to lift legal barriers in cases in which copyright and related rights could not be ascertained or in situations in which their owners were against it. However, the initiative of extending the provisions on reuse to cover all resources, including those under legal protection, requires the development of a new public policy on sharing public resources. The new reuse regulation does not address the above issues at the level of nation-states.

8.4 Social media regulation

The analysis of digital markets clearly defines the relationship between public authorities and the digital media environment. This is particularly true for an issue as troublesome as regulating digital content, including online media services. However, this view runs counter to the principle of the democratic will of the sovereign state, which pursues its own public interest, especially regarding cultural matters, when the equally fundamental

principles of subsidiarity and proportionality should be approached particularly seriously. Issues around regulating infrastructure and using tools for preventive content censorship are relevant, primarily because of the ever-changing shifts in the global position and role of market users. Technological advancements have increased the importance of infrastructure operators at the expense of content providers. This phenomenon is causing the entire digital world to be regulated at the technical and digital service access levels. This is why the service provider, or the digital service provider, is becoming so important. The examples of some countries prove the point that regulations adopted by public authorities in the field of digital media, whether more or less aggressive, are a means to strengthen the needs of the authorities, even in the most liberal areas. It is for this reason that public governance is today one of the main premises of public policy, including in the field of new technologies. As part of the planning process, the function of the public authority involves deciding on actions to achieve specific goals, and also redefining these goals by considering the requirements related to the development of modern technologies. All elements in advanced technologies, including software, digital services, and databases, exhibit the same characteristics, i.e., transfer rapidness (abrupt market changes in services, rapidly expanding technological innovations, especially in the context of network development, scientific research, etc.); globalisation (advanced technologies facilitating the global exchange of services in real time); entrepreneurship (the formation of cartels to conduct joint research for innovation, public-private partnerships); social participation (the development of innovative solutions by Internet users, the development of social media, crowdsourcing – the exchange of thoughts and concepts – all contributing to the growth in technology and innovations); convergence (combining multiple areas of human activity; technological convergence blurs the lines between individual fields in the legislative process, making it impossible to pinpoint threats and define liability through legislation, and to define the legal system alone), and result from freedom to exchange content. These characteristics also explain the need for changes and transformations related to establishing a new management system using legal instruments. Based on these important factors for developing new technologies, the case should be made to highlight the need to integrate the legal system in the most troublesome field of content regulation. Given the pace of technological and, in consequence, economic changes, this system must be characterised by flexibility, and the corresponding legal solutions should include universal standards allowing their application in different conditions and situations, depending on the nature of the digital content.

8.5 Legal security of digital content trading

Digital content has become a digital currency. This involves content created in the digital environment and that which has undergone digital conversion. It should be noted that this content may be subject to various kinds of protection. As regards determining the nature and legal status of digital content, on the one hand, a "test" is applied, which defines the following three elements: the data subject's activity in providing the data, the data subject's awareness of providing the data, and the data controller's activity in collecting

the data. On the other hand, to adequately protect user-provided content on the digital market, trading in such content should be based on three rules: the permission to use user-generated content, information obligations regarding content, and the right to withdraw and transfer such content from one platform to another. But, as in the case of personal data, the principle of exercising the right to transfer digital content, along with the right to erase such content (the demand to erase), should also apply. Legal transactions involving digital content should be based on transparent information for processing purposes, in order to control the principle of freedom of consent and the principle of purpose limitation. It is additionally essential to comply with the rule of transparency in business transactions on the ICT network. A key issue regarding the security of digital trading in digital content is determining the legal conditions under which "user-provided content" is protected in terms of copyright and the data subjects whose data is processed through various data-sharing platforms. This category of content may be subject to intellectual property protection within the framework, *inter alia*, of the Directive on Copyright and Related Rights in the Digital Single Market, as well as protection in terms of the subject-matter of e-services (e-commerce, in particular the proposed Digital Content Directive). The set of legal regulations referred to here is not uniform and is characterised by fragmentation in relation to specific sectors. For this reason, a cross-sectoral analysis appears indispensable to develop common definitions and options to manage user-provided private content on the digital market. The notion of user-provided digital content consists of several elements which should be taken into consideration in the context of the initiative in providing such content – the activities of the entity participating in its collection or processing and the extent of making use of such digital content by the platform user (which also includes taking into account third parties regarding the entity providing the data).

The convergence of digital media with traditional media has contributed to a special conflict which concerns defining the scope and level of the new regulations. This involves, in particular, digital content in which most issues touch on new media and new technologies (the protection of intellectual property, national identity, the right to privacy, and children and young people), as well as the economy (media market restrictions and the liability of digital service providers). New content management models (including online material) are emerging, accompanied by new rules for virtual organisation.

Evolution in communication technologies has materially changed the rules of the functioning of both individuals and societies. New multimedia platforms are being created to provide services electronically, which require modern technological solutions, usually financed by the private sector. An open and free global cyberspace allows cultural and experiential exchange across countries, societies, and individuals. It facilitates interaction and information sharing, leading to the spread of knowledge, experience, and technology. Freedom of speech and freedom of communication form the ideological grounds for such exchanges. The digital reality facilitates the performance of public tasks in a new social dimension (Chałubińska-Jentkiewicz, 2021:189).

The new technological order constitutes a premise and, simultaneously, the subject-matter of the discussed changes, which materially impact the regulatory area formed by digital media. Regulations applicable to this area of activities comprise four main aspects.

First, effective communication (i.e., freedom of speech as a fundamental right – Article 10 of the Convention for the Protection of Human Rights and Fundamental Freedoms, Article 54 of the Constitution of the Republic of Poland, and Article 1 of the Press Law Act). According to T. Garton Ash, a free press is a distinctive feature of a free country, while censorship is a characteristic feature of a dictatorship. "A democracy cannot long survive without the former, a dictatorship without the latter" (Garton Ash, 2018:295). (One should stress that the issue of restrictions on the freedom of speech in a democratic system concerns decisions to interfere with the content of religious organisations, the owners of "free" media, political factions, and other social groups which adopt a strictly defined way of thinking – which is supported by digital media – thus enclosing themselves within **an ideological bubble.** This not only applies to the receipt of certain information by network users but also to sharing digital content. Philosopher Onora O'Neill was right to note that our media "must not only be accessible to but also assessable" (O'Neill, 2011, as cited in Garton Ash, 2018:301). However, as stressed by T. Garton Ash, "By rights, the most effective constraint on the media should be us, the readers, viewers, and users" (Garton Ash, 2018:379).

Second, political and cultural diversity (i.e., worldview pluralism) – this issue is discussed in the context of the public interest in the widely understood media space, in consideration of the rules of public morality and public interest, provided that the existing regulations are of a protective nature (for instance, Article 30 of the Treaty on the Functioning of the European Union, Article 31 of the Constitution of the Republic of Poland, and Article 18 of the Act on Radio and Television Broadcasting). This should also apply to the new regulations still lacking in the field of new digital media, which appear indispensable because of the arrangements regarding liability for shared digital content. The worldview pluralism and the global exchange of thoughts foster creativity and form significant elements in the development of societies while also serving the purpose of consolidating their classic bases, identity, and cultural diversity, which should be considered in any elements subject to future regulation. It should be noted that major opponents of pluralism and diversity in digital media are their owners (vast Internet platforms or media corporations). According to A. J. Liebling,"The freedom of the press is guaranteed only to those who own one" (Liebling, 1975:32), considering that "what we have in a one-paper town is a privately owned public utility which is Constitutionally exempt from public regulation, which would be a violation of the freedom of the press".

Third, regulations justified by economic reasons – this mainly concerns the market economy on a uniform digital market. The need for such solutions arises from the principle of the uniform application of competition rules across the European Union. The

issue of selecting the forms of digital media market regulations, resulting from Poland's membership in the European Union, appears crucial in this context. Therefore, in the vast majority, these are EU regulations forming the basis for analysing documents, directives, and proposals defining a given regulatory area. According to the above-cited A. J. Liebling, "The function of the press in society is to inform, but its role in society is to make money" (Liebling, 1975:32). This rule can serve to define the purposes of the currently operating large Internet platforms.

On the one hand, they constitute an example of fragmentation and concentration which, to a great extent, results from binding self-regulatory principles. On the other hand, they are an efficient model of making money by exchanging user data, digital content, and databases without generating excessive costs still incurred by traditional media in connection with pluralism and objectivism. Digital platforms cannot be assumed to provide thorough information justifying their functioning. Driven by the rule that comments cost nothing while facts cost a lot, they have no intent to bear editorial responsibility for shared digital content, the verification of which is rather expensive.

Fourth, public service (i.e., the public interest and its objectives in the digital media sector), which in particular requires redefining objectives and orienting regulations towards ensuring the broadly-defined cybersecurity (including the need for protecting against disinformation, the violation of the right to privacy, hate speech, and content harming public morality, as well as safeguarding national identity, sovereignty, and the raison d'état of individual countries).

It is worth noting here that traditional media, in pursuing their public mission, begin to lose significance when faced with the omnipotence of social media, in which network users publicise content for other users without any in-depth analysis of what fulfilling public duties in media really is. Neither are such discussions conducted by those entities for which the media serve the purpose of implementing their diverse objectives, which are not related to any public interest.

The role of a nation-state – as borders no longer seem to matter in the global network context – is visible and fulfillable (at least to some extent) at each level, jointly forming the regulatory area of new digital media. The digital era has triggered the need to analyse the regulations applicable in this entire sphere. It appears necessary to modify the current forms of ***public authority*** as regards protecting public interest. This results from the fact that not all the executive instruments currently applied in the traditional approach to public administration duties (restrictions – registers, concessions, rules of territorial jurisdiction, and even the basic implementation of the legal regulations applicable to new digital media, e.g., cybercrime prosecution or editorial responsibility) are practicable in the new digital environment, which shapes not only a range of social behaviours and attitudes but also a new quality of the relationship between the state and individuals. Redefining public interest in this new and unregulated world should be coupled with

searching for new instruments to protect that interest and establishing new responsibility rules for shared digital content.

The business of digital content providers consists of making content available through information and communication systems. This category is highly diverse. It includes not only specialised institutions or entities but also end users. The latter group is particularly active due to the growing popularity of user-generated sites (or user-generated content). Due to the active form of their operations online, content providers bear direct liability for any breaches caused by such operations.

In Poland's legal system, content providers are directly liable for infringements of third-party rights. As noted by J. Barta and R. Markiewicz, controversies arose around attempts at qualifying the issue of making works available in computer networks (Barta, Markiewicz, 2001:228). Ultimately, this was qualified as a new field of commercialisation, i.e., making a work available in a manner that it could be accessed by anyone at any time and place they choose. This issue was highly relevant for ICT networks whose functioning was based on interactivity. As a result of digital processes, users can modify and share content without problems. The concept of *sui generis* protection of content producers' or providers' rights appears interesting. It was discussed at the *Association Littéraire et Artistique International* (ALAI) congress in 1996, with attempts to formulate a construct allowing producers to claim protection against third parties. Among others, consideration was given to affording them the status of moral rights or quasi-moral rights, with the caveat that they might not have limited the moral rights of content creators (Dietz, 1997; as cited in Gęsicka, 2014:290). According to J. Barta and R. Markiewicz, the construction of these rights is similar not to moral rights but to the author's economic rights (Barta, Markiewicz, 2001:228). It was this core objective, primarily economic, that these entities had in view, bringing these rights closer to related rights.

Table 1: Public interest in the new media v. New risks

Public-interest objectives in the media as presented to date	**New threats connected with the development of digital media**
protection of pluralism and opinion diversity	digital divide
protection of national and European culture from the domination of mass culture	new type of social exclusion
protection of children and young people	weakening of citizenship, cultural and national identity
protection of human dignity, no discrimination	cybersecurity
consumer protection	weakening of the right to privacy and lack of anonymity
	infringement of the ownership right (copyrights to digital content, databases)
	loss of data confidentiality
	information war and disinformation

When talking about the changes being brought about by new technologies in the digital content-sharing environment, we must remember that this development requires an interdisciplinary approach, combining the knowledge and experience of experts in the fields of economics, sociology, technology, the media, political science, psychology, culture, and security science. Today's living conditions largely depend on the state of the information and communication technologies functioning in a given country. We are currently witnessing radical changes both in how societies function and in the global economy because of the expected spread of innovative information and communication solutions. Freedom of speech and freedom of communication form the ideological grounds supporting such exchanges. Thanks to new mass media technologies (ICT networks, the Internet) – a subject that has been explored particularly extensively – entirely new and previously unknown approaches to family, professional, and public life have emerged. Along with the development of digital technology and social changes, including those associated with forming the so-called digital democracy, new fields of human action have emerged, commonly described as the ICT network environment, and more broadly understood as cyberspace. They affect all aspects of human life. The same is true for social and economic relations, and the state-individual accord, which includes exercising fundamental human rights. The digital revolution we are witnessing, including, in particular, automated data-processing technologies, which affect human-related decision-making processes, takes us back to the questions about human rights and freedoms. Whereas regulatory restrictions previously applied to the relations between the state and the individual, current normatively enshrined steps taken by public authorities are becoming a means to protect the rights and freedoms of individuals first and only then to ensure public security, order, and morality, in a world driven by technology used for a wide range of purposes, except that behind each *technē*, even the most automated, there

is a person. Aristotle (384–322 B.C.) maintained that a distinction should be made between *technē* (practical skills, art) and *epistēmē* (scientific insights, knowledge) (Aristotle, 2005:114, translated by Piotrowicz). According to this philosopher, knowledge forms include all sciences, while art and practical skills represent inferior occupation types associated with craftspeople and slaves. According to this concept, *technē* is an obstacle to practising virtue in the souls and minds of the free. This historical approach related to contempt for *technē*, which is usually not associated with such notions as ethics, public morality, or personal interests, continues to be relevant, especially in the context of the right to privacy, the protection of personal interests, and moral standards (Chałubińska-Jentkiewicz, Karpiuk 2015: 6). In the field of modern technology, there is an ongoing conflict between them and so-called sensitive interests. An open and free global cyberspace allows cultural and experiential exchange across countries, societies, and individuals, facilitating interaction and digital content sharing, and leading to the exchange of knowledge and experiences. Hence, it can be said that digital content sharing facilitates the exchange of technology, thereby driving innovation. The development of new technologies and the associated processes of social changes require a new regulatory approach and a redefinition of public-interest objectives and public-authority responsibilities in the process of regulating the areas which are relevant to the core aspects of the functioning of the individual – citizens, markets, and states. Convergence processes occurring in so far differently understood regulatory areas contribute to the rise of a special type of conflict as to the scope and level of new regulations. Difficulties arise in specific globalisation conditions, exterritorial digital services, and due to the absence of universal state jurisdiction and sovereignty rules.

In the modern-day cyberspace realm where individuals function, it seems necessary to establish norms and, before this, rules and values to apply as standards in the real world. Freedom in the online environment also requires security and protection and, consequently, regulatory restrictions. However, due to the nature of cyberspace, new needs must be considered. This includes establishing new values, also those specific to that environment. This is particularly evident when most of our activities have moved to the online realm due to the COVID-19 pandemic. In particular, this applies to defining the roles of Internet users and the rules of liability for online activities. Yet, this is only one piece of the highly complex issue of advanced-technology development in the context of the legislative process, affecting almost every state, society and the weakest of all links – the individual.

The state must gradually limit the scope of its governance function in favour of shaping development, standardisation, and mediation strategies and mechanisms. An important part of this function is to make projections about forecasts. This requires an extensive analysis of local and global considerations, economic, social, and political needs and interests, and the possibility of meeting individual needs. A diagnosis and strategy would help to formulate the appropriate regulatory policy, which is closely linked to the realm of governance. The sphere of development governance differs from the other three areas

of public administration functioning in that it is oriented more towards the future functioning of the state. It is for this reason that governance is now one of the central premises of public policy, including new technologies. To take these measures, public authorities need norms. These revolutionary changes involve state government (including the entire e-government area), chiefly because previous state government and governance methods will prove ineffective in a society where information has become the main instrument and digital content – the primary product. Modern technologies have created administrative convergence – a process whereby new, common administrative solutions are developed to replace traditional administrative divisions. These areas are usually defined at the EU level. They are divided along the lines of new threats to the rights and freedoms of individuals – and to the European Economic Area. One of the key regulatory objectives in the legislative process is to guarantee cybersecurity, which requires the accessibility and integrity of networks and infrastructures, and also, most importantly, the confidentiality of digital data processed within them and their ownership protection, as well as their security against illegal content. This means that protecting the fundamental rights and freedoms of individuals sets the bounds within which each legislative process should take place. This also applies to drafting legislation due to the development of new media, including social media. Freedom of speech and the right to communication are not absolute values, and, as such, they represent no obstacles to regulations geared towards the public interest, security, public order, public morality, and the rights and freedoms of other individuals (Article 31(3) of the Constitution of the Republic of Poland).

All elements in advanced technologies, including software, services, databases, and equipment, exhibit the same characteristics, i.e., rapidity (abrupt market changes in digital services, rapidly expanding technological innovations, especially in the context of Internet development, data processing (automatic profiling), etc.); globalisation (advanced technologies facilitating the global exchange of digital services on the digital market in real time); entrepreneurship (the formation of consortia to conduct joint research for innovation, public-private partnerships); social participation (the development of innovative solutions by Internet users, the growth of social media, crowdsourcing – the exchange of thoughts and concepts – all contributing to the growth in technology and innovations); convergence (combining multiple areas of human activity; technological convergence blurs the lines between individual fields of the legislative process, making it impossible to pinpoint threats and define liability through legislation, and to define the legal system alone). These characteristics explain the need for changes and transformations related to establishing a new system for digital content management using legal instruments. Based on these important factors in developing new technologies, a case should be made to highlight the need to integrate the legal system in cyberspace.

The assessment of the digital markets in the above-mentioned scenarios – for instance, based on Freedom on the Net reports – clearly defines the relationship between public authorities and the digital media environment. This is particularly true for an issue as

troublesome as regulating electronic media content, including online digital services. However, this view runs counter to the principle of the democratic will of a sovereign state, which pursues its own public interest when the equally fundamental principles of subsidiarity and proportionality should be approached particularly seriously. Issues around regulating the operations of digital service providers (as in the case of editorial responsibility and publisher's liability) and using tools for preventive censorship in digital content are relevant primarily because of the ever-changing shifts in the global position and roles of market users. Technological advancements have increased the importance of infrastructure operators – or, more specifically, content distribution platforms – at the expense of the providers of the same content. This development leads to the entire digital world being regulated at the technical and network organisation access levels. Some examples of the proposals for digital market regulations discussed in this treatise prove that regulations adopted by public authorities in cyberspace, whether aggressive or not, are ways to strengthen authorities' needs, even in the most liberal areas. This is particularly relevant as the digital content-sharing world is in growing need of regulations.

References:

Angelopoulos, C. & Smet, S. (2016) Notice-and-Fair-Balance: How to Reach a Compromise between Fundamental Rights in European Intermediary Liability, *Journal of Media Law*, 8(2), http://dx.doi.org/10.2139/ssrn.2944917.

Barta, J. & Markiewicz, R. (2001) *Internet a prawo* (Warszawa: Wydawnictwo Universitas).

Chałubińska-Jentkiewicz, K. & Karpiuk, M. (2015) *Prawo nowych technologii. Wybrane zagadnienia* (Warszawa: WoltersKluwer).

Chałubińska-Jentkiewicz, K. (2021) *Prawne granice dezinformacji w środkach masowego przekazu* (Toruń: Wydawnictwo Adam Marszałek).

Cooter, R. & Ulen, T. (2011) *Law & Economics* (Boston: Addison-Wesley).

Dietz, A. (1997) General Report, In: Dellebeke, M. (ed.) *Copyright in Cyberspace* (Amsterdam: Otto Cramwinckel).

Frosio, G. & Mendis, S. (2020) Monitoring and Filtering: European Reform or Global Trend?, In: Frosio, G. (ed) *The Oxford Handbook of Online Intermediary Liability* (Oxford: Oxford University Press).

Galloway, S. (2018) *Wielka czwórka. Ukryte DNA: Amazon, Apple, Facebook i Google* (Poznań: Wydawnictwo Rebis).

Garton Ash, T. (2018) *Wolne słowo, Dziesięć zasad dla połączonego świata* (Kraków: Wydawnictwo Znak).

Geere, D. (2012) *How deep packet inspection works*, available at: https://www.wired.co.uk/article/how-deep-packet-inspection-works (October 4, 2024).

Gęsicka, D. K. (2014) *Wyłączenie odpowiedzialności cywilnoprawnej dostawców usług sieciowych za treści użytkowników* (Warszawa: WoltersKluwer).

Horten, M. (2016) *The Closing of the Net* (London: Polity Press).

Liebling, A. J. (1975) *Wartość liczbowa odnośnie do rywalizujących gazet codziennych New Yorker* (New York: New Yorker).

Lucas, E. (2017) *Oswoić cyberświat. Tożsamość, zaufanie i bezpieczeństwo w internecie* (Warszawa: Kurhaus Publishing).

Ombelet, P. J., Kuczerawy, A. & Valcke, P. (2016) *Employing Robot Journalists: Legal Implications, Considerations and Recommendations* (Montreal: The Web Conference), pp. 731-736.

Rossi, A. (2020) *After Truth: Disinformation and the Cost of Fake News*, available at: https://www.youtube.com (February 17, 2021).

Saba Bebawi, S. & Bossio, D. (2014) *Social Media and the Politics of Reportage: The 'Arab Spring'* (London: Palgrave Macmillan London), pp.1-8.

Based on the #Digital2020 report, available at: https://mobirank.pl/2020/02/23/digital-mobile-i-socialmedia-w-polsce-w-styczniu-2020-roku (July 20, 2020).

Urban, J., Karaganis, J. & Schofield, B. (2016) Notice and Takedown in Everyday Practice, *UC Berkeley Public Law Research Paper*, No. 2755628, https://dx.doi.org./10.2139/ssrn.275562.

Chapter II

Disinformation and Cyberterrorism in Light of the Standards of the Council of Europe

KATARZYNA CHAŁUBIŃSKA-JENTKIEWICZ & MONIKA NOWIKOWSKA

Abstract We live in an age where the internet is the main source of information. We can find reliable facts there, but it is also easy to come across misrepresentations, half-truths or news stories that are only intended to sow panic. This is how disinformation works. In the age of the internet, manipulating information has become a powerful tool. Disinformation in the new media is entirely intentional and deliberate. It involves the transmission of false or manipulated information that causes the recipient to be misled. Disinformation creates an image of the world that is inconsistent with reality. It leads to erroneous decisions and actions and creates a false view of a particular piece of information. Disinformation can influence election results, shape public behaviour and affect the mood of a country. Disinformation on the internet has become one of the biggest threats to the digital space. Today, it is not limited to individual states, but affects institutions at the international level.

Also linked to technological advances and the development of the global internet is the phenomenon of cyber-terrorism. Cyber-terrorism is a combination of classic terrorist activities and the use of the latest ICT devices. States are becoming increasingly aware of the threats emanating from the network and are taking up the fight against them in order to protect the most important elements of critical infrastructure that guarantee the smooth functioning of a country. States need to be ready at both a legal and practical level for the occurrence of a cyber-terrorist attack in order to be able to effectively repel and defend themselves. This article

CORRESPONDENCE ADDRESS: Katarzyna Chałubińska-Jentkiewicz, Ph.D., Associate Professor, Kozminski University, College of Law, ul. Jagiellonska 57/59, 03-301 Warszawa, Poland, e-mail: kasiachalubinska@gmail.com, ORCID: 0000-0003-0188-5704. Monika Nowikowska, Ph.D., Assistant Professor, War Studies University in Warsaw, Faculty of Law and Administration, Aleja Generała Antoniego Chruściela „Montera" 103, 00-910 Warszawa, Poland, e-mail: m.nowikowska@akademia.mil.pl, ORCID: 0000-0001-5166-8375.

https://doi.org/10.4335/2024.2.2 ISBN 978-961-7124-25-5 (PDF)
Available online at http://www.lex-localis.press.

examines the issue of Disinformation and Cyberterrorism in Light of the Standards of the Council of Europe.

Keywords: • cyberterroris • cecurity threats • human rights • journalist • disinformation

1 Introductory remarks

Disinformation constitutes a serious security threat for contemporary democratic societies – states, international organisations, and individuals. It should be stressed that this phenomenon is becoming one of the most significant and complex challenges of the 21st century. As an element of measures related to threats, disinformation is a phenomenon that resembles terrorist actions. As regards the notion of terrorism, despite comprehensive studies of the subject, no widely acceptable definition has been developed since 1937 when the League of Nations prepared the first draft Convention on the Prevention and Punishment of Terrorism. This results from the fact that there are fundamental differences in opinions, attitudes, and interests, arising from historical, cultural, and religious conditions. The situation is similar to the definition of disinformation. Currently, in the context of the rapidly developing new threats to security and public order, characteristic of the convergence era, activities based on information technology are becoming fundamental. This refers to the preparation and implementation of individual undertakings as well as to organising and financing decentralised networked structures. The shift of the paradigm, as part of which the traditional forms of actions, including acts of terrorism, are disappearing, gives rise to the fact that the sphere of new threats is being identified, including interrelations between terrorism and digital media, or digital services in general, considered to be the key accelerator of changes to the global information system, which in turn constitutes the foundation on which contemporary societies are shaped. It is currently possible to speak about the emergence of a clear cultural pattern which brings together the spheres of telecommunications, information technology and media, and forms an intricate system, a peculiar multi-communication environment spanning multiple levels, distribution platforms and types, including printed media, linear and non-linear audio-visual services and their Internet forms, so-called digital media (Chałubińska-Jentkiewicz, 2023a: 228).

Speaking of disinformation as an act of terrorism in the theatre of contemporary times, or referring to digital media as the oxygen thanks to which it can assume completely new forms, has become the canon of defining mutual relationships between them. However, research into this sphere has not revealed any such straightforward links. Scholars do not offer clarity as to a comprehensive theory indicating major trends in respect of the relationships between disinformation as an act of cyberterrorism and digital media (Nacos, 2009: 4–5). On the one hand, terrorists use the existing social communication media as an effective distribution platform, create them, or are active users of new communication and information services, particularly social media. On the other hand, digital media constitute an intriguing area of wholly new threats, from the perspective of the mission undertaken by the media, the nature of marketing activities and the commercial approach. Terrorists strive to attract global public attention to the objectives and causes for which they organise and conduct their operations. Carrying out attacks, often targeting unspecified, anonymous and numerous victims, is aimed at evoking fear, and the widespread dissemination of information about such attacks may favour its

intensification, at times compelling public authorities to make decisions that the terrorists expect. It is similar with disinformation. The regular use of digital media for disinformation purposes, as in the case of terrorism, might also result in their legitimisation, according to the agenda setting concept. As Ch. de Franco aptly notes, "the narratives produced by the media, especially those constructed around one or more images, do create a reality effect which impacts not only the public at large but also policymakers. Those narratives constitute a mediated reality which interferes with the policymaking process because they affect the mental image of a given issue through which policymakers interact and based on which they take decisions" (Franco, 2012:47).

Digital media, which are currently functioning in an environment characterised by unprecedented competitiveness, where information flow, in addition to its cultural dimension, is becoming an important economic sector, are searching for pieces of news that are highly attractive to the audience. It seems that the pressure for fast and topical messages favours their tabloidisation, and dramatic and bloody images are somewhat consistent with the expected pattern, so one can speak about their overrepresentation in social communication media. Disinformation may be built around such stories.

2 Cyberterrorism

It should be stressed that a lot of online information might affect the types of targets and weapons selected by terrorists and their operational methods. Cyberterrorism consists of using information technologies, i.e., computers, software, telecommunications devices, and the Internet, to reach the goals that a given group has set. As B. Hołyst aptly noted, "just like multiple corporations use the Internet for making their activities more effective and flexible, terrorists leverage the power of technology (IT) to develop new operational doctrines and organisational forms" (Hołyst, 2011:63). The emergence of terrorist groups linked in a network constitutes a part of a concept called netwar. Cyberterrorism involves the disruption or destruction of opponents' information systems and the seizure of their strategic data. Terrorist organisations using the web are characterised by informal communication depending on their needs and cross-border reach, i.e., moving beyond state borders, dispersion and mutual trust, with no hierarchical bureaucratic structure (Chałubińska-Jentkiewicz, Nowikowska, 2020: 305).

While cyberterrorism is defined as a phenomenon characterised by a high degree of abstraction, the progressing development of information technologies allows the statement that the risk of a terrorist cyberattack is increasing. The actuality of the deployment of such cyberattacks stems from the fact that terrorists use the Internet to plan and conduct physical attacks, to spread ideologies, to manipulate public opinion and the media, to recruit and train new terrorists, to acquire and build up funds, to obtain information on potential targets, to control the operations being conducted, or to gain access to confidential information constituting a secret of various types (Smarzewski 2017: 66).

From the perspective of cyberterrorism, new communication methods reduce transmission time, which allows online participants-terrorists to communicate despite being dispersed, as well as to reduce communication costs and to extend the scope and comprehensiveness of information. In addition to network forms of terrorist organisations, IT also contributes to improving the collection and analysis of materials as part of terrorist intelligence activities, consisting of the search for attack targets via the Internet. The above conditions facilitate the deployment of various types of offensive informational operations, such as propaganda campaigns (recruitment of members, acquisition of funds, and public outreach) – attacks against virtual targets (electronic attacks, computer system choking, sending of unsolicited e-mail at a mass scale, web bugs) – used for physical damage.

3 The Council of Europe's standards on combating disinformation

Since the beginning of its existence, the Council of Europe has taken up the topic of terrorism on multiple occasions, generally placing the issue in the sphere of cooperation within the justice system in criminal matters. It should be stressed that the perspective was unchanged, determined by human rights and the need to protect them. This meant balancing initiatives taken to maintain fundamental rights and freedoms as the key values defining the shape of its axiological system. The Council of Europe adopted two conventions: the Convention on the Prevention of Terrorism of 16 May 2005 (OJ EU L 159/3) and the Convention on Laundering, Search, Seizure, and Confiscation of the Proceeds from Crime and on the Financing of Terrorism done in Warsaw on 16 May 2005 (Journal of Polish Law 2008, No. 165, item 1028). Article 1(1) of the Convention on the Prevention of Terrorism includes a definition of a terrorist offence, which means any of the offences within the scope of, and as defined in, one of the treaties listed in the Appendix to the Convention. Under the Convention, the Parties are obliged to establish public provocation to commit a terrorist offence, recruitment and training for terrorism as criminal offences (Articles 5–7). Taking the above solutions into account, a general definition of an act of terrorism can be adopted, according to which it is any offence committed by individuals or groups resorting to violence or threatening to use violence against a country, its institutions, its population in general or specific individuals which, being motivated by separatist aspirations, extremist ideological conceptions, fanaticism or irrational and subjective factors, is intended to create a climate of terror among official authorities, specific individuals or groups in society, or the general public. Therefore, it spans across the multitude of contemporary forms of the phenomenon being discussed, from organised, international group "undertakings" to single acts committed on the territory of a given state by individuals, motivated by irrational or subjective factors, as stated above (Chałubińska-Jentkiewicz, 2023a: 230).

It is worth noting here that disinformation activities constitute a vital part of terrorism and cyberterrorism. From the perspective of Article 10 of the Convention for the Protection

of Human Rights and Fundamental Freedoms (ECHR), referring to the freedom of expression, the Council of Europe defined the scope of such relationship in its Declaration of 2 March 2005 on freedom of expression and information in the media, in the context of the fight against terrorism. The authors stressed the negative impact of the phenomenon of terrorism on human rights and referred to the need to achieve unity between the Member States of the Council of Europe to unequivocally condemn all acts of terrorism as criminal and unjustifiable, threatening and destabilising social life, wherever and by whoever committed. In this context, governments face a challenge to balance the need to uphold the freedom of expression, as the foundation of democratic and pluralistic societies, and the assurance of security (Chałubińka-Jentkiewicz, Nowikowska 2022: 20). It was asserted that the free and unhindered dissemination of information and ideas is one of the most effective means of promoting understanding and tolerance, which can help prevent or combat terrorism. This also applies to the phenomenon of disinformation. As per Article 10(1) of the ECHR, everyone has the right to freedom of expression, including the freedom to hold opinions and to receive and impart information and ideas without interference by public authority and regardless of frontiers. The freedom of expression exercised by the media, including digital media, is not absolute and unlimited, as it carries responsibility and the resulting obligations. It should be stressed that the freedom of expression may be subject to restrictions which, however, do not go beyond the boundaries set by the provisions of Paragraph 2 of the Article being discussed, construed in line with the case law of the European Court of Human Rights, unifying its interpretation.

From the perspective of disinformation related to the phenomenon of cyberterrorism, it is possible to speak about restricting the freedom of speech based on such criteria as public safety and, given the increasing threat of cybercrime, preventing the disclosure of information protected by law (Chałubińska-Jentkiewicz, Nowikowska 2022: 115). Measures derogating from the obligations under the ECHR may be taken in times of war or other public emergencies threatening the nation's life. However, they should not be contradictory to other obligations under international law (Article 15(1)). The states have been obliged to make every effort to refrain from adopting measures threatening the freedom of the media, constituting one of the pillars of democratic societies, particularly exploited by disinformation actors. Therefore, every instance of restricting the freedom of expression must be subject to formalities prescribed by law and necessary in a democratic society (Chałubińska-Jentkiewicz, 2023a: 230).

As already mentioned, disinformation relies on digital democracy. It should be noted that, in the context of limiting the freedom of expression, the Committee of Ministers called on public authorities in Member States not to introduce any new restrictions unless they are strictly necessary and proportionate in a democratic society and subject to examining carefully whether existing laws or other measures (hard or soft) are not already sufficient, and to refrain from adopting measures equating media reporting on the phenomena of terrorism with support for terrorism. Moreover, as regards the issues being discussed,

Member States were obliged to ensure access by journalists to information regularly updated, in particular by appointing spokespersons and organising press conferences, with due respect for human dignity and subject to the right to respect for private life. Journalists should also have access to, follow, and report on judicial proceedings and the judgements referring to persons who are the subject of anti-terrorist judicial proceedings, with due respect for the presumption of innocence (Article 6(2) of the ECHR). It was stressed that any potential restrictions meeting the aforementioned requirements may be based on the criterion of the so-called good of the justice system, in special circumstances and to the extent strictly necessary in the opinion of the court, where publicity would prejudice the interests of justice (Nowikowska 2023: 113). The press may be excluded from all or part of a trial, for example, in the interests of national security (Article 6(1)). It is also vital to respect media independence and the right not to disclose sources of information and also to refrain from any pressure on the media, etc.

Firm suggestions are also addressed to the media, particularly to journalists, whilst they are made from the perspective of their responsibility for the contents being disseminated. This means that they should not support terrorist organisations or the operations they conduct by, for instance, offering a platform to terrorists to present their objectives and ideas, giving them disproportionate attention, adding to the feeling of fear in society, or even unintentionally serving as a vehicle for violence through the expression of racist or xenophobic feelings or hatred. However, this should not be coupled with self-censorship, as the effect of this would be to deprive the public of necessary and desired information, including expert opinions and results of consultations. In addition, the media should not disseminate information in situations when such actions would jeopardise the safety of persons and the due conduct of anti-terrorist operations or judicial investigations of terrorism. Finally, the media should not abstain from respect for the right to dignity and private life, particularly concerning the victims of terrorist attacks and their families (Article 8 of the ECHR). They should also adhere to the rule of the presumption of innocence regarding potential perpetrators, taking into account the distinction between suspected or convicted terrorists.

It should also be asserted that the media should bear in mind the positive role they can play in preventing hate speech, promoting mutual understanding, and creating an atmosphere of tolerance. Press representatives are also encouraged to hold training courses on the broadly understood theme of terrorism, from its historical, cultural, religious and geopolitical aspects to practical issues related to improving their safety, and to invite journalists to follow these courses. Following the above recommendations, the media should adopt self-regulatory measures or adapt the existing alternatives to laws to effectively respond to the ethical issues raised by media reporting on terrorism.

The Committee of Ministers of the Council of Europe monitors the implementation of the above recommendations and suggestions by the governments of Member States, in particular in the legal field, considering the issues from the perspective of standards

related to fighting terrorism and their effect on the freedom of expression in the media. The assumptions put forward in the Declaration are confirmed in Recommendation 1706 (2005) Media and Terrorism. The Parliamentary Assembly of the Council of Europe, referring to previous documents concerning the fight against terrorism in the context of human rights, including the freedom of speech, stressed the significance of the rights of democratic societies to be informed about matters of public concern. This also includes such acts as threats and terrorist attacks or the response by the state and international organisations to these threats and acts. At the same time, the Parliamentary Assembly indicated that terrorist acts were intended to create terror, fear or chaos among the public. The effect of such acts depends largely on how they are reported in the media. Messages that are disseminated at a global level and repeated multiple times are dramatised and sensationalist. They often result in distorting and exaggerating the real issues out of proportion. The public and the media must be aware that perpetrators intentionally utilise such acts to have the strongest possible impact.

However, the above considerations do not change the fact that, subject to the right to privacy, it is essential to inform the public about terrorist acts. In specified cases, properly disseminated information might contribute to forming adequate political responses. The Assembly also recommended that Member States take account of this recommendation in their national work and hold a debate on this issue in their respective national parliaments, inform the public and the media regularly about government strategies and actions, and inform, upon their request, media about the specific situation to avoid journalists investigating terrorism being unnecessarily exposed to dangers. It is also important to cooperate with law enforcement authorities to prevent the dissemination of illegal messages and images by terrorists on the Internet. Member States should place special emphasis on abstaining from prohibiting or even restricting unduly the dissemination of information and opinions in the media, as well as on the reaction by state authorities to terrorist acts and threats under the pretext of fighting terrorism. Of course, the recommendations refer to information about terrorist activities, but they also include a significant context related to disinformation.

The media are encouraged to develop, through their professional organisations, a code of conduct for journalists, photographers, and other professionals dealing with the subject-matter to keep the public informed about terrorism issues, in line with the highest professional standards. This is a crucial matter. It also includes the need to organise training courses for media professionals to increase their awareness of the sensitive nature of media reports on the issues in question. In particular, emphasis was placed on the cooperation between individual media entities and their professional organisations to avoid a race for sensationalist news and shocking images which violate the privacy and human dignity of victims or increase the negative effect of such acts on the public, which is what terrorists expect. It is also important to avoid aggravating fear and the societal tensions underlying terrorism, and to refrain from disseminating any hate speech by offering terrorists a platform for presenting their news, views, and opinions. As already

mentioned, the Council of Europe also issued Guidelines of the Committee of Ministers of the Council of Europe on protecting freedom of expression and information in times of crisis, adopted on 26 September 2007. It is generally based on the principle of freedom of expression, treated as the basis for the functioning of contemporary democratic societies and the personal development of every human being. The protection covers not only desirable and non-offensive information and ideas but also messages that are shocking and not widely acceptable (Nowikowska, 2020: 54). This stems from such values as pluralism and tolerance, which are essential to the Council of Europe. In particular, the protection should include broadly understood freedom of speech in matters of public concern, artistic expression or commercial communication. Therefore, Member States that impose restrictions must substantiate a strong societal need, making such interference indispensable, legitimised based on publicly available domestic laws, and falling within European standards. The restrictions must be proportional to their objectives, and it should be noted that stricter limitations, for example in relation to penal sanctions, will require stronger justification. At the same time, the Council of Europe's standards do not authorise absolute and unlimited access to classified government information. The above approach should be applied in times of crisis, where authorities are especially tempted to impose restrictions on society. The Council of Europe condemns all violent acts, including the killings of media professionals, while stressing the need for dialogue between governments, media professionals and civil society to guarantee freedom of expression. Journalists play a crucial role in times of crisis by providing accurate, timely and comprehensive information. They also have the capacity to foster a culture of tolerance and understanding between different social groups. According to the provisions of Section I, state authorities are free to adopt their definition of crisis, whilst the Council of Europe has provided a relevant framework, giving examples of such situations in the form of a non-exhaustive list. Terrorist attacks are one of them. The term *times of crisis* is not associated with an officially and legally introduced state of war or other emergencies but it generally refers to the factual circumstances. Similarly, a comprehensive definition of "media professionals" is proposed to ensure the widest possible protection to all persons working in the information flow sector. In Section II, an obligation was placed on Member States to ensure, to the fullest extent, national and foreign media professionals' safety, provided that measures taken to this end must not be used by Member States as a pretext to unnecessarily limit the rights of media professionals, such as their freedom of movement and access to information or areas affected by the crisis. Restrictions may be applied only when absolutely necessary. Authorities should also provide regular information to all media professionals covering the events equally through various channels, e.g., press conferences. If possible, they should set up information centres. The above postulates result from the threats that journalists encounter while on a mission to inform the public about crises (Chałubińska-Jentkiewicz, 2023a: 236–237).

Military and civilian agencies in charge of managing crises are also expected to take practical steps to promote the understanding of crises, including the ones related to

terrorism or cyberterrorism, in cooperation with media professionals dealing with these issues. Employers should strive for the best conceivable protection of their media staff on dangerous missions, including by providing safety equipment, comprehensive counselling (from legal to psychological), and life and health insurance. Furthermore, journalism schools and professional media associations are encouraged to provide specialised safety training for media professionals. Another significant requirement is that member states should protect the right of journalists not to disclose their sources of information, especially information referring to the identity of informants, as the foundation of personal safety and the control function that media professionals play (Nowikowska, 2023:106). Moreover, they should not misuse in libel and defamation legislation against media professionals, and thus limit their freedom of expression. Times of crisis do not entitle states to restrict the freedom of expression of the media beyond the limitations allowed by Article 10(2) of the ECHR, especially in matters of key importance to the public. When imposing potential restrictions in the event of, e.g., incitement to violence or public disorder, such terms should be adequately and clearly defined. It is also necessary to consider that the media might contribute to resolving crises as, for instance, public service media might be a vital factor for social integration between various groups. In times of crisis, Member States' maintenance of a favourable environment for freedom of expression and independent media, in line with the standards set by the Council of Europe, should also include the possibility of criminal or administrative liability for those public officials who try to manipulate public opinion by exploiting its special vulnerability. This might take place in specific matters concerning the examination of whether certain information or documents should be revealed to journalists, and the final decisions in this respect (Chałubińska, 2023a:238).

In times of crisis, such as terrorist attacks, the process is, to a large extent, affected by the inclination to disclose partial, manipulated, or even false data. In the discussed situations, the media also have a special responsibility as they are expected to adhere to the highest professional standards, including ethical ones. In such circumstances, the regular provision of factual, accurate, timely and comprehensive information to the public can play a major part in awareness-raising and calming down public sentiments. In transmitting such information, as regards its content, form and context, the media should be attentive to the rights of other people, their distinct sensitivities, and their possible feelings of uncertainty and fear.

Digital media are developing separate guidelines, partly in fear of the regulatory measures that public authorities might but, generally speaking, they are not convinced about such solutions, as they require the widest possible extent of freedom and operational flexibility. In this respect, cooperation is needed between self-regulatory bodies at the national, regional and European levels, coupled with support from state authorities and other stakeholders engaged in these issues.

It should be added that the amendment to the Audiovisual Media Services Directive introduced significant modifications in this sphere. As part of implementing its provisions in the Polish legal system, obligations concerning digital content for video-sharing platforms were introduced (bearing in mind that, according to the definition of such platform, it also includes a place where users share other content, not just video files). Furthermore, the Digital Services Act (the DSA) refers to all online platforms and other online service providers operating in the EU, including marketplaces, e.g., Amazon, social media and search engines. The obligations laid down in the said Regulation depend on the size of a given enterprise – the larger the entity, the more extensive the list of obligations. The categories of very large online platforms (VLOPs) and very large online search engines (VLOSEs) include companies which have the average monthly number of active recipients of the service in the Union equal to or higher than 45 million. Major American platforms (i.e., Google, YouTube, Amazon, Apple, Meta) fall within these categories. Enterprises defined as VLOP and VLOSE will have to continuously analyse and mitigate so-called systemic risks, such as the dissemination of illegal or harmful content (e.g., disinformation) or manipulation of users' behaviour. VLOPs will also be obliged to provide (national and Union-level) supervisory authorities and researchers with access to the data and algorithms that would allow a detailed assessment. The DSA also provided a crisis response mechanism. As part of the mechanism, if an event posing a threat to public safety or health occurs (such as the Russian invasion of Ukraine or the COVID-19 pandemic), the European Commission may oblige VLOPs to adopt specific measures, for instance, to remove for three months selected contents that spread harmful disinformation or accounts of users who incite dangerous behaviours. Comprehensive and constructive dialogue between government authorities, the media and other domestic entities interested in combating disinformation and the establishment of a platform for debates favour the assurance of freedom of expression in times of crisis. It should be added that Directive 2018/1810 does not provide grounds for sanctioning user activities beyond the right to restrict access to contents that violate the provisions of the Directive. The only sanctions imposed on users include blocking such content and limiting the possibility of publishing new content. In addition, users may be subject to liability on general terms if the contents they publish infringe the provisions of other legal acts (for example, if they contain child pornography or incite terrorism) (Chałubińska-Jentkiewicz, 2023a: 239).

Cooperation at an international level, particularly with the Council of Europe and other organisations, facilitating information exchange and monitoring possible violations effectively, is also desirable. Non-governmental organisations have the potential to contribute to the safeguarding of freedom of expression and information by monitoring infringement of the freedom of speech in various ways, such as maintaining helplines for consultation, reporting harassment of journalists and other alleged violations targeting the media and their mission. Such entities should also cooperate in offering comprehensive support and training to media professionals. The addressees of the guidelines should include Member States, media organisations, and other interested civil society entities.

Nonetheless, unlike other spheres of communication operations where responsibility is distributed in similar proportions, in the case of disinformation strictly related to the category of security, the burden of implementing the Council of Europe's standards should be essentially imposed on domestic public authorities. As for normative standards referring to human rights, states were left with substantial flexibility in assuring public safety and order, consisting of the possibility to introduce restrictions on the freedom of expression under the ECHR, particularly taking into account the public safety criterion. Although they are obliged to refrain from introducing new restrictions other than the ones that are strictly necessary and proportional in a democratic society, and only where existing legal instruments and other alternative measures are insufficient, and although the criteria for the establishment of restrictions on freedom of expression are listed on a *numerus clausus* basis in Article (2) of the ECHR, such criteria are defined in detail at a national level. It should also be noted that the temptation to put in place restrictions towards cross-border activities and media operating at a global scale intensifies in the circumstances of a crisis, where terrorist acts become more severe and violent. What is more, such restrictions are more willingly tolerated or even approved by the public in such circumstances.

In general, the Council of Europe's standards concerning the protection of the freedom of expression do not require any changes. However, their implementation at the Member State level might give rise to certain doubts. Regarding restrictions, public authorities may adopt extremely diverse approaches, ranging from a *laissez-faire* policy to censorship. Self-regulation or co-regulation is a potential third option. Moreover, the objectives of governments and the media are not always convergent. While the mass media usually strive for complete independence, effectiveness, also in commercial terms, and safety of its operations, governments expect that they should support the objectives, strategies and, at times, even specific operations conducted as part of counteracting the practices discussed in this paper. They also expect that the perpetrators of terrorist acts are presented as criminals whose conduct cannot be justified in any way.

The media are, on the one hand, seen as the pillar of rights and freedoms, including freedom of speech, and as a factor facilitating the spread of disinformation, on the other hand. This gives rise to the yet unsolved dilemma of whether, when, and to what extent public authorities may introduce restrictions on access to information. The states may refer to issues related to the criteria for restrictions in a more precise way, also at the interpretation level, using the case law of the European Court of Human Rights, which is obvious. Guidelines should be developed in close cooperation with the media. Concerning self-regulatory measures being applied by the media in the sphere of disinformation, if they prove to be ineffective, the concept of co-regulation should be considered in the scope in question (preferably in the initial approach formula). Summing up, it can be stated that refraining from censorship is a crucial principle resulting from the Council of Europe's standards. In the context of disinformation, this measure should not be excluded, for example, if a given (online) medium is directly controlled by hostile

foreign services or if illegal content needs to be removed or blocked. Features of terrorist activities may be noted in disinformation campaigns, which brings the phenomenon closer to cyberterrorism acts, i.e., acts of aggression in cyberspace. Cyberterrorism is a multidimensional phenomenon covering financial resources, state-of-the-art technologies, and broadly understood logistics. The power of cyberterrorism, as a certain branch of terrorism, stems from the fact that one person having specialist knowledge and equipped with basic computer devices can paralyse air traffic, affect the transmission of electricity or cause a failure of banking systems, robbing ordinary citizens, institutions or even state enterprises of their funds, as well as influence human behaviour, stance, and emotions. The combat against cyberterrorism is very problematic and laborious due to the vastness of cyberspace and the challenges related to locating perpetrators. A large proportion of such offenders are still unattainable to law enforcement authorities. One of the ways to fight cyberterrorism is to cut off funds by eliminating financing sources. Other methods consist of developing a stable strategy model that would mark out shared activities in combating cyberterrorists and establishing international organisations to combat or mitigate cyberterrorism and to eliminate disruptions in state critical infrastructure. It is worth noting that, according to D.E. Denning, cyberattacks motivated by political objectives may be a manifestation of cyberterrorism. It is important to assert, however, that a situation must occur where not only the legal and economic order is disrupted, which gives rise to considerable loss whose dimensions are becoming purely material and physical, affecting people (Denning, 2002: 79). As can be noticed, to a large extent, terrorists increasingly often use non-conventional weapons and less complex modes of their operations. Disinformation measures may be characterised by the properties of cyberterrorism. Attracting attention is the basis for existence in digital media. This, in turn, is necessary for the so-called agenda setting to work. It is a concept according to which information in the media is treated as significant by the audience. Disinformation evokes and acts on fear to reach political transformations. Undoubtedly, we are dealing here with a form of psychological warfare. Vivid examples of how the atmosphere of fear can be built effectively across society include the disinformation activities related to COVID-19 and 5G. The theory concerning the origin of the coronavirus was the greatest fake news of all time. According to the thesis, it was a biological weapon created to destroy a competitive economy. Another fake news that added to the atmosphere of fear was the link between COVID-19 and 5G. Fake information was published on social media saying that the emitted electromagnetic radiation would accelerate the spread of the virus. The effects of this absurd theory were real, as mobile network towers were burnt in numerous cities across Europe.

The objective of disinformation actors using the media was to disseminate their convictions, ideologies, and motives. The Internet allows them to achieve the objective to a greater extent than the traditional media. Terrorists are becoming the authors of their image and are not dependent on journalists assuming their role. This allows them to demonstrate their operations not as barbaric acts but as an uneven battle between an oppressed group of partisans and world powers. Numerous videos posted on such

websites as YouTube may serve as an example. It is not the intention of disinformation actors to be seen positively. What they mostly want is to present their ideologies and demands. It can be described as propaganda through action. A part of their message is lost in information noise and is misrepresented or forgotten by recipients. All that is remembered is the slogan, for example, fighting against Ukrainian fascists.

Current or potential members or enthusiasts of a given theory belong to a significant target group that various actors try to reach through the media. Enthusiasts are a vital element which every organisation needs, as they are the ones who support disinformation activities. Such assistance may be effective as they are not directly related to the informational message. Thus, they are not responsible for its content. A clear example of how the atmosphere of support for a given theory can be built effectively was disinformation related to the anti-vaccine movement. In the analysis by the Academic Centre for Cybersecurity Policy (*Analysis Concerning the Impact of the Social Phenomenon of Anti-Waxxers on the Security of the Polish State,* performed by Inserq sp. z o.o. for the Academic Centre for Cybersecurity Policy, dated 12 December 2021), to fulfil research objectives, two types of objects were used: a Twitter account and threads. Each of the objects offers different analysis possibilities. The study included the identification of specific Twitter accounts which had a specified influence on Polish information space and generated the greatest quantities of digital content, and specific information about COVID-19 and vaccines that was further shared on Twitter. In early 2020, i.e., at the beginning of the pandemic, when the SARS-CoV-2 virus started spreading across Europe, it was noted that the main axis of public interest included the informational content produced by major opinion-forming media in Poland, e.g., Fakty TVN, Polsat News, TVP Info, as well as media centres related to medicine, for example, the Ministry of Health, MedOnet and other outlets. According to the collected data, in the initial phase of the pandemic in Poland and Europe, there were no groups strictly and explicitly negating the existence of the SARS-CoV-2 virus and the COVID-19 disease. Instead, the dominating place was taken by the emotions of fear, the demonstration of a strong will to obtain the greatest possible number of pieces of information on the threat, mockery of the circumstances and the fear, and opinions that the virus is far away (in the Far East), so there is no need to discuss it. Accounts that negated the existence of the virus in general, its mortality rates and the devastated health of infected patients began appearing in the media space around March and April 2020, and the highest surge of such accounts, including the most popular ones followed by tens of thousands of users, was recorded around mid-2020. Moving from negating the existence of the virus to negating the existence, necessity and effectiveness of using vaccines against SARS-CoV-2 was a natural continuation of the trend (Chałubińska-Jentkiewicz, 2023a: 245–246).

4 The phenomenon of acceptance and justification of disinformation

As regards terrorism, an interview with a terrorist, conducted by media representatives, can be seen as a kind of legitimising such acts. For instance, the TVN24 channel decided to take such a step in interviewing Ali Ağca. Another dangerous trend that can be observed in the media is attempting to understand terrorists' motives, resulting in the unintentional justification of their conduct. Experts, often invited to television studios, try to refer to cultural, social, economic, or psychological considerations. Undoubtedly, the most important factor for media operations in the free digital single market are numbers of viewers, listeners or readers who decide to use a given communication medium. The proceeds to the budget of a given media institution (mostly from advertisers) depend on them. Therefore, potential recipients must be provided with a product that is "attractive" enough for them.

Media theory authors indicated properties that a given event should have to become a valuable media product for the audience. These include 1) timeliness – an event should be "fresh", preferably published nearly real-time; 2) intensity – the spectacular and intense nature of the event has a positive impact on its media value; 3) unambiguity – an event should be easy to assess by most recipients; 4) importance – an event should be important in the sense of its impact on society; 5) conformity – understood as meeting the audience's predictions and expectations, which may be based on stereotypes; 6) surprise – the extraordinary nature of a given event; 7) continuity – an event should last for an extended time; 8) references to prominent individuals or major international relations actors – events that affect the most important entities are interesting; 9) complementarity – the possibility to link a given event with specific individuals or past events; 10) negativism – negative events are more spectacular than the positive ones (Chałubińska-Jentkiewicz, 2023a: 246–247).

Given the above list of characteristics of newsworthy events that are attractive to the audience, it can be concluded that disinformation messages meet most of the criteria. The media are not only used by disinformation actors but also leverage the newsworthiness of disinformation-related events for their own purposes. For instance, false information was disseminated in the media concerning a deadly game called the Blue Whale Challenge. The game was allegedly to cause the death of over 130 teenagers. According to the thesis, teenagers aged between 14 and 17 were to complete challenges assigned by their mentor. The objective was to strive for their death. The matter gained publicity when *The Sun*, a British tabloid, wrote about it. The information was further copied by several Polish websites, after which facts were mixed with fiction and passed on by nearly all mainstream media. "Hyperbolisation" is one of the characteristics of the Internet. And this property was used in the Blue Whale story. So far, no reliable sources or tangible evidence have been identified to prove that such danger occurred.

"For commercial media, breaking news, often the most tragic, is – horribile dictu – a blessing. This is their logic, and ethics will not be able to do much about it (...)". Therefore, the media are showing blood, sensationalism, and human drama. A colourful tabloid and a reputable opinion-forming newspaper will approach such events differently (but none would disregard it). Various ways of approaching a given theme are called formatting.

The following piece of news presented in various formats may serve as an example:

1. Informational format (agency-style): "Four people were killed and 33 were injured after a bomb exploded in a café in Paris on Thursday morning".
2. Sensationalist format (in a reputable newspaper): "A bomb thrown by a terrorist in a busy Paris café lethally wounded four people and left 33 others covered in blood".
3. A story format (a piece of news in a tabloid newspaper, illustrated with a huge photograph showing scattered remains): "A newlywed couple on their honeymoon died on Thursday when a bomb destroyed a café in Paris. The young wife and husband, who had got married a day before, were among the four killed and 33 injured in a bomb explosion".
4. Educational format (a commentary in a serious newspaper): "The bomb attack in a Paris café on Thursday seems to herald a new wave of violence inflicted by Islamic fundamentalists outraged by French foreign policy in the Middle East".

Media experts have noted that, currently, we are dealing with a shift towards reporting on events in the tabloid story format. This is because the mass audience expects reader-friendly information which is spectacular at the same time. The process has also spread across the informational activities of online users. The development of digital democracy has contributed to changes in the media market in the economic and organisational context, and in the information sphere, taking into account the quality and significance of information itself. We are dealing with media power, which can be defined in several ways. N. Couldry and J. Curran define it as a label for the net result of organising a society's resources so that the media sector has significant independent bargaining power over and against other key sectors (big business, political elites, cultural elites, and so on) (Couldry, Curran, 2003:39). The power defines most relationships that are formed around the media and are practised at multiple levels, individual and collective players, organisations, institutions and networks of connections. Since power is practised at various levels of media activities, relationships of power are multidimensional and complex, especially in light of the emergence of new forms of media practice, such as networked journalism. Networks are defined as "complex structures of communication constructed around a set of goals that simultaneously ensure unity of purpose and execution flexibility by their adaptability to the operating environment. They are programmed and self-configurable at the same time. Their goals and operating procedures are programmed in social and organisational networks, by social actors. Their structure evolves according to the capacity of the network to self-configure in an endless search for more efficient networking arrangements. This definition by Castells suggests that

networks are ever changing and evolving towards a higher degree of efficiency, which in turn means a higher degree of power" (Bebawi, Bossio 2014: 125). In the 19th century, groups opposing state authorities created their own means of communication, aware of the role the media plays. An anarchist newspaper, The Truth, published in the USA, may serve as a good example. Its slogan said: "The Truth costs 2 cents, and dynamite is 40 cents a pound. Buy them: read the paper, use the dynamite". Vladimir Lenin also spoke about the establishment of media independent of state authorities. Revolutionary press, both legal and illegal, was needed to "agitate, propagate, and organise". Also, Carlos Marighella argued that, despite reports of the activities of revolutionaries/guerrillas/terrorists in official media, they should establish their means of communication (Chałubińska-Jentkiewicz, 2023a: 249).

5 Concluding remarks

The Internet has become a medium of fundamental importance. Thanks to the Internet and advanced technologies, the arsenal of terrorist communication methods has been extended by multimedia materials, audio and video recordings, blogs, and other websites, utilising numerous interactive tools, such as fora, discussion lists, chats or messaging apps. Furthermore, the properties of websites are favourable to the activities of groupings opposing state authorities. Internet advantages that terrorist organisations may benefit from include a) easy access, b) limited state control, c) the possibility to reach a wide audience, d) anonymous activities, e) the speed of information transfer, f) low cost, g) media convergence (multimedia), and h) the possibility to influence traditional media that often use the Internet to search for information. Relying on these properties of the Internet, terrorist groups use the web to conduct propaganda and publicity operations of the group, gain supporters, communicate within their internal structures, recruit and mobilise terrorists, or acquire funds.

Based on the above deliberations, a conclusion can be drawn that online operations create possibilities to reach a wider public than traditional media. Moreover, the Internet has become a medium resembling a worldwide press agency. Information posted online by terrorists are likely to be used by traditional media and websites. Videos of hostage executions by terrorists or terrorist group leaders' appeals may serve as an example here. Such types of news are published on websites related to terrorists and then spread across the Internet, reaching television, radio and press.

Secondly, it can be stated that disinformation has become a form of entertainment whose advantage over other forms consists in its sensationalist nature. Disinformation actors provide attractive topics to the media which use the opportunity meticulously to generate profit. Reports on activities in Ukraine can be cited as examples of how appropriately selected tactics for presenting acts of terrorists or military operations can affect the market position of a given medium. Hybrid operations are a constant part of Russian strategies towards other countries. Russia relies on various narratives for its propaganda, operations

of influence and psychological operations, although they may be grouped by their mutual features (Nowikowska, 2022: 164–165). The attempts to prove the alleged hatred in the mutual relationships between Poles and Ukrainians was a fairly popular trend in disinformation activities. An example of this is a false piece of information that appeared in the Ukrainian and Russian-speaking media sphere in September 2019. It was about an alleged murder of a Ukrainian soldier committed by a Polish soldier. The murderer was to be a Polish instructor from the Joint Multinational Training Group Ukraine, and the offence was said to have been committed in Javoriv near Lviv. The place was not randomly selected. The training ground was crucial not only for security building in this part of Europe but also for Polish and Ukrainian relations in the military sphere. At the time, it was a ground intended for the operations of Joint Multinational Training Group Ukraine, established by Poland, the USA, Canada, Denmark and Lithuania, to support the Ukrainian army to allow them to reach NATO compliance standards (Gliwa, 2022). It is worth noting that, by analysing Polish information space, we can clearly see that military operations were conducted concurrently with activities in the information sphere. Disinformation following the Russian invasion and the refugee crisis that it triggered had two directions. One of them was intended for the Polish information space (addressed to the Polish society), and the other one, referring to Poland, was destined for the global market. Given the second path, we can speak about the attempt to discredit, in the eyes of the international public, the work of Polish soldiers, border guards, and individuals selflessly helping the Ukrainians fleeing war.

Thirdly, online activities seem the best way to gain supporters and potential recruits. Any person keen on the ideology and operational methods of terrorists will search for the information they are interested in on the Internet. Being aware of the fact, terrorists publish a lot of material glorifying their attitudes. Thus, many young and frustrated people might get fascinated by radical views and, in extreme situations, even become new attackers. Information terrorism is a phenomenon which is developing and spreading actively, posing a threat to the entire global community.

Notwithstanding their motives, terrorists' overall objective is to attract the attention of public opinion and to intimidate a large number of people. The media plays a key role in terrorist organisations. Terrorists' strategy assumes, *inter alia*, making as many people worldwide as possible aware of such brutal incidents. Terrorists have effectively used the mechanisms of media influence on the audience for a long time. That is why they plan their attacks in a way that allows them to attract media attention and to place information on a given event on top of the daily information agenda. Information weapons include information resources which are strategically designed or built to conduct information warfare, to cause damage, confusion or inconvenience, or to carry out any other malicious activities. Information terrorism is characterised not only by cyberspace but also by manipulating and falsifying information, and, in some cases, also by creating false facts, as a result of which disinformation occurs to intimidate and evoke paranoid thoughts among the targeted population (Chałubińska-Jentkiewicz, 2023a: 254).

The characteristics of information terrorism include:

- organised violence, a specific type of psychological terror;
- dissemination via the media;
- psychological impact on a wide population;
- attention;
- intimidation and deprivation of the population;
- the surprise effect;
- public and ostentatious nature of operations (Chałubińska-Jentkiewicz, 2023a: 254–255).

Information terrorism may be additionally divided into: a) information and psychological terrorism, or media terrorism (controlling the media to spread disinformation, demonstrating the power of terrorist organisations to destabilise societies), and b) information and technological terrorism or cyberterrorism (damage to a specified part or the whole of the opponent's information environment).

Social media generally serve two basic functions to terrorists: the information and propaganda function and the tactical and operational function. According to a report prepared by experts from the United Nations Office on Drugs and Crime (UNODC), it is possible to list several key areas of Internet use by terrorist organisations: 1. spreading online propaganda, including the recruitment of new members and incitement to terrorist attacks, 2. financing, 3. training, 4. planning terrorist attacks, including the preparation and use of encrypted communications and use of open source intelligence; 5. executing attacks, 6. cyberterrorism. Considering the meaning of the notion of cyberterrorism, it should be stressed that legal commentators have aptly noted that cyberterrorism is something more than just a prefix added to standard terrorist activity (Smarzewski, 2013: 184). According to the definition by K. Liedel, cyberterrorism is a politically inspired attack or a threat of attack against computer information networks or systems, aimed to destroy infrastructure, intimidate governments and individual citizens or impose far-reaching political and social objectives upon them (Liedel, 2006: 36). D. Jagiełło uses a different definition of cyberterrorism, stating that it includes politically or militarily inspired attacks or a threat of attack against information and communication (ICT) systems and networks or collecting data to paralyse or severely damage state critical infrastructure, intimidate or impose far-reaching political and military actions on governments or communities, as well as the intentional use of ICT networks and the Internet by terrorist organisations, national liberation movements and insurgent movements to paralyse national critical infrastructure or to intimidate or impose specified conduct on governments or the population (Jagiełło, 2013: 12). A different approach was suggested by J.A. Lewis who states that cyberterrorism is the use of computer network tools to shut down critical national infrastructures (such as energy, transportation) or to coerce or intimidate a government or civilian population (Lewis, cited in Siegel, Worrall

2014: 638). In this sense, common features of disinformation and cyberterrorism can be observed (Siegel 2012: 385).

Finally, it should be stressed that the World Wide Web, one of the benefits of the digital era, has become a weapon of those trying to combat harmful phenomena and disinformation. The combat against disinformation in cyberspace may become one of the most significant challenges contemporary legislators will need to face. It is worth stressing here that counteracting disinformation should not consist of mass control and censorship of the Internet because the only winners, in this case, would be the enemies of one of the most important individual liberties, i.e., freedom of expression. The procedure for removing and blocking content might become one of the most significant solutions concerning disinformation. According to the currently applicable provisions of the Radio and Television Broadcasting Act, if user-published contents violate the applicable legal regulations, which includes disinformation, hate speech, contents containing aggression, or rules (which users are obliged to comply with under the Act on the Provision of Service by Electronic Means), the platform provider will be authorised to demand that such user remove the said infringements. If the contents are not returned to a legitimate state (through flagging or removal, depending on the type of violation), the platform provider will have the right to block access to such contents to other users. The contents will not be removed from a platform, and only the users who have published the contents will have access to them. After the contents are blocked, they will not be available to the general audience of the contents presented on the platform. Besides, in the event of further infringements by a given user, the platform provider can temporarily block the relevant account (to temporarily block the publication of new content). In the most serious situations, where the user concerned publishes contents that incite terrorism or include child pornography, the platform provider will be able to permanently prohibit such publications, for instance, by liquidating the account. Technology development is a challenge for future legislators and regulators. The responsibility for digital content is becoming the domain of intermediaries.

The most recent law governing this sphere is Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services and amending Directive 2000/31/EC (Digital Services Act) (OJ EU L. 277/1 of 27.10.2022). The DSA retains the responsibility rules applicable to service providers and intermediaries specified in Directive 2000/31/EC on Electronic Commerce, considered the foundation of the digital economy. In the DSA, the term “illegal content” was not defined in detail. As per Article 3(h) of the DSA, “illegal content” means any information that, in itself or in relation to an activity, including the sale of products or the provision of services, is not in compliance with Union law or the law of any Member State which complies with Union law, irrespective of the precise subject matter or nature of that law. This means that the identification of illegal content will be determined by the system of values in place in a given Member State (Chałubińska-Jentkiewicz, 2023b: 193). According to the intentions of EU legislators, the DSA will guarantee clear criteria for

removing online content, and ensure an effective complaint and redress mechanism in the event of blocking user content and its publication. New obligations arising from the DSA, for instance, those imposed on the providers of online platforms, include the assurance of transparent recommender systems and online advertising, the need to ensure the traceability of business users, the provision of services taking into account the fundamental rights, including the freedom of expression, taking care of appropriate measures to protect against misuse (mechanisms for users to signal such contents, and in the event of platforms mechanisms for cooperation with "trusted flaggers", or the establishments of points of contact aimed to ensure direct communication with Member States' authorities, the Commission, the European Board for Digital Services, and recipients of the services. The objective of the regulation is to protect citizens' rights and prevent disinformation.

New technologies are making us all smarter. Should we be concerned about the linking of existing values and the ever-present domination of technology? In the literature on the subject, it is indicated that, by 2045, humans will have multiplied their intelligence by a billion by connecting their cerebral cortex wirelessly with a new synthetic cortex in the cloud. Questions related to the security of such development and exploitation of this sphere remain unanswered. They are mainly related to transforming citizens into e-citizens, the divergent interests of market and political stakeholders, and the political arena. The most critical issue to resolve is who decides what is wrong and right, i.e., what is legal and why (Kerikma, Rull 2016: 13–14).

Perhaps mediation will become a vital part of combating disinformation, like the procedure for removing and blocking content in media laws. R. Hill identifies several elements of effective negotiations, such as approaches facilitating the achievement of a common position. He describes it as the power-negotiating tactic. The five pillars of the tactic have been described below:

1. "Don't react: go to the balcony". The author warns against excessively emotional reactions that might lead to confrontation. Sometimes, it is better not to react by expressing rigid and extreme positions but to step back and let things cool down before negotiations are resumed.
2. "Don't argue: step to their side". The author suggests not to argue but "turn" to the opposing side. Confronting what seems to be an unreasonable demand from one of the parties, one should not react by restating an extreme position. Instead, we should acknowledge the points both parties agree on and restate calmly our requirements. It is important to overcome suspicion and mistrust.
3. "Don't reject: reframe". In confronting an unacceptable request, it is better not to reject it immediately but to ask why the other party is making it and find ways to restate the problem so both parties can benefit from continuing their negotiations.
4. "Don't push: build them a golden bridge". While approaching understanding in a delicate matter, it is better not to push approval too intensely. Instead, we should find ways to evoke a sense that a shared position has been worked out.

5. "Don't escalate: use power to educate". If there is a threat of rejecting a compromise proposal, which could end negotiations, it is better not to escalate the problem through pressure. A more effective solution is to calmly indicate the consequences of the lack of consent and inform the other party about the advantages of the compromise and the problems that might arise if it is not reached (Hill, 2014: 148–150).

References:

Bebawi, S. & Bossio, D. (2014) *Social Media and the Politics of Reportage The 'Arab Spring'* (New York: Palgrave Macmillan).

Chałubińska-Jentkiewicz, K. (2023a) *Prawne granice dezinformacji w środkach społecznego przekazu. Między wolnością a bezpieczeństwem* (Toruń: Wydawnictwo Adam Marszałek).

Chałubińska-Jentkiewicz, K. (2023b) *Cyberodpowiedzialność. Wstęp do prawa cyberbezpieczeństwa* (Toruń: Wydawnictwo Adam Marszałek).

Chałubińska-Jentkiewicz, K. & Nowikowska, M. (2020) *Bezpieczeństwo, tożsamość, prywatność – Aspekty prawne* (Warszawa: C.H. Beck).

Chałubińska-Jentkiewicz, K. & Nowikowska, M. (2022) *Prawo mediów* (Warszawa: C.H. Beck).

Couldry, N. & Curran, J. (2003) *Contesting media power: alternative media in a networked world* (London: Rowman & Littlefield Publishers).

Denning, D. (2002) *Wojna informacyjna i bezpieczeństwo informacji* (Warszawa: WNT).

Franco, Ch. (2012) *Media Power and the Transformation of War* (London: Palgrave Macmillan).

Gliwa, S. (2022) Polak winny morderstwa ukraińskiego żołnierza? To fake news, *CyberDefence24.pl*, available at: https://cyberdefence24.pl/fake-news/polak-winny-morderstwa-ukrainskiego-zolnierza-to-fake-news (March 20, 2022).

Hill, R. (2014) *The New International Telecommunication Regulations and the Internet* (Heidelberg: Springer-Verlag GmbH Berlin).

Hołyst, B. (2011) *Terroryzm* (Warszawa: LexisNexis).

Jagiełło, D. (2013) Cyberterroryzm, *Edukacja Prawnicza*, (5), pp. 10-14.

Kerikma, T. & Rull, A. (2016) *The Future of Law and eTechnologies* (Switzerland: Springer International Publishing).

Liedel, K. (2006) *Bezpieczeństwo informacyjne w dobie terrorystycznych i innych zagrożeń bezpieczeństwa narodowego* (Toruń: Wydawnictwo Adam Marszałek).

Nacos, B.L. (2009) Revisiting the Contagion Hypothesis: Terrorism, News Coverage and Copycat Attacks, *Perspectives on Terrorism*, 3(3), pp. 3-13.

Nowikowska, M. (2023) Tajemnica dziennikarska o dobro procesu karnego, *Palestra*, 5, pp. 101-115.

Nowikowska, M. (2020) *Granice dozwolonej krytyki prasowej działalności osób pełniących funkcje publiczne* (Warszawa: C.H. Beck).

Nowikowska, M. (2022) SYOPS as an element of information warfare, In: Chałubińska-Jentkiewicz, K. & Evsyukowa, O. (eds.) *Information disinformation cybersecurity* (Toruń: Wydawnictwo Adam Marszałek), pp. 163-170.

Siegel, L.J. (2012) *Criminology* (Belmont: Cengage Learning).

Siegel, L.J. & Worrall, J.L. (2014) *Introduction to Criminal Justice* (Belmont: Cengage Learning).

Smarzewski, M. (2013) Bezpieczeństwo państwa jako przedmiot ochrony niektórych przestępstw popełnianych za pośrednictwem sieci teleinformatycznej, In: Dziemianko, Z. & Kijas, A. (eds.)

Bezpieczeństwo współczesnego świata. Edukacja i komunikowanie (Poznań: Wydawnictwo Wyższej Szkoły Handlu i Usług), pp. 173-192.

Smarzewski, M. (2017) Cyberterroryzm a cyberprzestępstwa o charakterze terrorystycznym, *Ius Novum*, (1), pp. 64-74.

Chapter III

Regulatory Dilemmas Around Social Media

JĘDRZEJ SKRZYPCZAK

Abstract This chapter aims to answer whether there is a need, or even a necessity, for legal regulation of social media today. It is also necessary to analyse by what methods (whether 'hard' regulation is necessary or whether self-regulatory solutions are sufficient) and at what level (national, regional, international) such regulation should be introduced in order to, on the one hand, ensure the effectiveness of such solutions, given the specificity of social media functioning, and, on the other hand, respect freedom of speech. While today there is no doubt that some regulation of social media is necessary, one should call for it to be done with great caution. Furthermore, this is true both in terms of the scope of such regulation and the method and reach. In considering the need for appropriate regulation in this area, it is argued that the temptation to regulate the activities of such platforms may lead to a restriction of freedom of expression, with the measures adopted serving to censorship and restrict public debate.

Keywords: • social-media • regulation of social media • self-regulation

CORRESPONDENCE ADDRESS: Jędrzej Skrzypczak, Ph.D., Professor, Head of Department of Media Systems and Media Law, Adam Mickiewicz University in Poznan, Faculty of Political Science and Journalism, Uniwersytetu Poznańskiego 5, 61-614 Poznan, Poland, e-mail: jedrzej.skrzypczak@amu.edu.pl, ORCID: 0000-0002-5906-3802.

https://doi.org/10.4335/2024.2.3 ISBN 978-961-7124-25-5 (PDF)
Available online at http://www.lex-localis.press.

1 Introductory remarks

This article will examine whether there is a need or necessity to regulate how social media operates. If such a need or necessity indeed exists, it would be warranted to consider the methods ("hard" regulations or self-regulatory solutions) with which to fulfil them and the levels (national, regional or international) at which they should be implemented, to make sure the solutions are effective in the complex social-media environment. The analysis was primarily based on comparative, inductive and deductive methods, and on legal exegesis.

2 Literature review and theoretical framework

The emergence and turbulent development of social media have been a unique phenomenon of the digital era (Kreft 2016: 17). These media have come to exemplify "demassified" means of communication (Toffler 2002: 447; Dziemba 2009: 53–61; Grzesik-Robak 2009: 27–35; Młynarska-Sobaczewska, Preisner 2008: 113–127; Palmer, Eriksen 1999: 32) and blurred boundaries between the producers (senders) and consumers (recipients, audiences) (Veltman 2006: 3–47; Kowalski 2003: 23–30; Krzysztofek 2007: 223–224), allowing users to both receive and create content. Essentially, social media operates on the principle of users' sharing content(Zafarani et al. 2014: 1). Communications become personalised as everyone can receive content in a one-to-many or one-to-one model. Legal authors and commentators offer many definitions of the social media phenomenon. For example, according to one interpretation, the term refers to "a group of web applications based on the ideological and technological foundations of Web 2.0, designed to facilitate the creation and exchange of user-created content (Kaplan, Haenlein 2010: 59–68). As of January 2022, about 62.5% of the world's population had access to the Internet, and 58.4% used social media (Digital 2022). Indeed, being online today is largely about using social media.

Whilst there are many types of social media, they all – regardless of their operational form, platform, concept and rules – have several commonalities. First, they represent a type of personalised communication. Second, content creation on social media happens when one user communicates with another. Third, they nonetheless manage to attract large audiences. For instance, all social media allow using bots, algorithms and fake accounts to distribute content and create – in a premeditated and purposeful manner – an illusion that certain subjects or persons arouse great public interest. Providing a pathway for content to circulate freely around the world, these networks can facilitate major manipulation and disinformation campaigns. Hence, there is a paradox in which social media provide immense opportunities for freedom of speech, all the while having the potential to jeopardise it. It is important to stress that social media platforms and administrators differ from traditional media. A significant difference is that they neither create any content nor interact with its authors. They only provide the digital space to share information and opinions. This leads some to argue that such platforms should not be accountable for the content published in discussions and user interactions. One may

wonder what it is about social media that makes it a vehicle for hate speech and a trigger of infodemics. Is it only about the illusion of anonymity? Or perhaps the ease of content creation is the culprit?

Whatever the answer, the key lies in determining whether social media is a private or public space. On the one hand, the biggest media platforms (such as Facebook, Instagram and X) are mostly US-based corporations (TikTok being the exception) with specific commercial objectives. In this regard, they should effectively enjoy economic freedom, with minimum state intervention in the form of a general legal framework for operating a business. On the other hand, they often become the main platform for public debates on critical issues such as elections. Since social media can win elections and influence public opinion, it may be necessary to regulate them (Patterson, 2020; Kumm, 2023;Stahl (2020)**.**

3 Freedom of speech protection standards from the analogue era

With the development of human rights standards, the 20th century – the era of analogue media – saw the establishment of guarantees for the freedom of the press and speech. These were enshrined in various documents adopted by the UN, UNESCO, the Conference on Security and Co-operation in Europe and the Council of Europe, and in the European Court of Human Rights case law (Gardocki 1993: 111). States followed suit by implementing similar safeguards in their respective legal systems. However, most of these regulations came into use in the era of analogue media, usually long before the dawn of the Internet and social media. Hence, there are legitimate doubts about their applicability in the digital age.

The aforementioned international regulations guarantee universal freedom of opinion and expression. They describe it as the freedom to hold undistorted opinions, and to seek, receive, and impart information and many ideas through whatever medium, regardless of frontiers. This means that international law affords the right to hold and – notably – share opinions. Freedom of speech is a basic human right. It is essential to the functioning of democratic societies and vital for the growth and development of states and individuals. Freedom of speech should encompass not only neutral but also derogatory statements. It should also be noted that these norms apply to various media regardless of the technology they rely on to distribute their content (Skrzypczak 2019: 81–92). Freedom of speech is, however, not an absolute right and may be subject to restrictions in specific circumstances. It is also important to remember that, under Article 19(3) of the International Covenant on Civil and Political Rights, the exercise of the rights provided for in this paragraph carries special duties and responsibilities. It may, therefore, be subject to certain restrictions, but these shall only be such as are provided by law and are necessary: (a) For respect of the rights or reputations of others; (b) For the protection of national security or of public order (*ordre public*), or of public health or morals. Similarly, Article 10(2) of the European Convention on Human Rights stipulates that "the exercise of these freedoms, since it carries with it duties and responsibilities, may be subject to

such formalities, conditions, restrictions or penalties as are prescribed by law and are necessary in a democratic society, in the interests of national security, territorial integrity or public safety, for the prevention of disorder or crime, for the protection of health or morals, for the protection of the reputation or rights of others, for preventing the disclosure of information received in confidence, or for maintaining the authority and impartiality of the judiciary".

4 Do we need to regulate social media?

For some time now, an important and considerable debate has been taking place on whether it is necessary to introduce regulations on how social media operates (Tully 2014: 53–172; Paslawsky 2017: 1486; Khan 2021; Tan 2021; Barrett 2021; Brannon 2021; Kayode-Adedeji, Oyero, Aririguzoh, 2018: 393–439; Scaife 2021). More recently, the discussion became more heated, mainly due to the 2021 events in the US. To recall, on 6 January 2021, supporters of the outgoing US President Donald Trump stormed the Capitol. Twitter blocked two tweets by Trump on the grounds of them being "potentially misleading" (Wall, Mooppan et al. 2021; Varis 2021; Garcia, Hoffmeister 2017). He responded by accusing Twitter of meddling with the campaign. This did not solve the problem. On the contrary – the platform blocked Trump's official account, boasting a considerably large following of 88 million users (Kreft 2021: 13). Facebook and Instagram followed suit (Ohlheiser, Guo 2021). In addition to banning Trump's account, Twitter took many other measures, including blocking over 70,000 accounts linked to the QAnon conspiracy theory, while Facebook started blocking posts containing the "Stop the Steal" slogan. Other platforms implemented different solutions to remove content and adopted internal reforms. For instance, YouTube targeted Trump's account by removing videos instigating violence and imposed a seven-day ban on uploading new content to Trump's account. Meanwhile, Stripe stopped processing payments for Trump's campaign website. A fierce debate ensued on whether digital platforms had the right to censor public debate without any judicial authorisation. The main concern was when left to their own devices, they could effectively influence election results (Palmer 2021). Were we, in fact, dealing with a "privatisation" of censorship in these cases? Twitter's then CEO Jack Dorsay admitted, at one point, that he was not proud of blocking Trump's account but that this was a good decision for the platform (Diaz; Kreft 2021: 16). The overall response to the situation was negative, with social platforms facing serious accusations of being a breeding ground for "extremism, disinformation and sociopaths managing profit-driven algorithms – the viruses behind the Capitol epidemic we have witnessed" (Kreft 2021: 16; Galloway 2021). For many, Facebook's and Twitter's bans were long overdue but there was also a large group condemning these steps as freedom of speech violations. These people were asking when it was warranted for these essentially private entities to "de-platform" individuals – especially well-known public figures such as Trump – and how they should go about it. It was the platforms that took the blocking measures in the case in question. Some believe it was sufficient evidence that self-regulation in their industry was adequate (Garcia, Hoffmeister 2017). However, there is a view that social media have transformed from the growth factor they once were into what is now a public-

order disruptor fuelling the "us versus them" sentiment. According to some opinions, it was not a coincidence that the political leaders who thrived on social media were those pursuing a divisive agenda. It is stressed that one of these platforms – Facebook – is currently the biggest news distributor in the world although the news they provide is principally anger- and hatred-driven lies. This is attributable to the fact that such messages catch on and spread faster than neutrally dull facts. The "a lie told a million times becomes a fact" adage seems to apply here. There is no truth without facts. Without truth, there can be no trust. Without them, democracy as we know it is "dead" (Ressa 2020).

Similarly, a 2021 UN study showed that online hate speech, especially on social media, was a growing phenomenon worldwide.

It is important to note that major US corporations mostly own social media. In the United States of America, there is a long tradition – dating back to the First Amendment – of protecting free speech, even if considered offensive. Conversely, many European democracies approach freedom of speech differently and have no qualms about legislating bans on hate speech.

As a result, demands for regulating this aspect of social media have been increasingly common and forceful (Ressa 2020:17; Fox 2021). The task, however, would be laden with numerous challenges and dilemmas. A popular view among legal authors and commentators is that in the analogue era, responsibility for guaranteeing freedom of speech rested with two types of entities: states and international organisations. In today's digital age, a third actor comes into play – private corporations which own global communication platforms. In the analogue era, states and international organisations played a key role. Now, the balance of responsibility is shifting to corporations holding influence over content published on their global platforms. These include mainly content aggregators, such as social media platforms, which can limit some content and activities if they deviate from their internal rules and – as shown by practice – marketing strategies (Papernik 2022). At the same time, proponents of regulating this area have expressed their concern that regulators might feel tempted to restrict freedom of speech and to legislate measures that effectively censor and stifle public debate. It is important to remember here that most of the major players in the social media realm are private entities – corporations formed under US law but operating on a global scale. This raises serious questions about whether, at what level – international, regional or national – and how to regulate these platforms effectively. Another dilemma relates to what type of regulation would be the most appropriate – hard law or soft law combined with self-regulatory measures. An alternative option would be to refrain from legislative steps and instead focus on promoting safe use practices on social media (Balkin, 2018).

5 Global regulatory framework

It is important to note here that, in September 2018, the International Commission on Information and Democracy was appointed to define the rules governing the global information and communication space, guided by the principle that it is "a common good of humanity". French President Emmanuel Macron introduced this initiative during the G7 Summit in Biarritz. Later, at the meeting of the Alliance for Multilateralism as part of the 2019 UN General Assembly, he put it forward as the Partnership for Information and Democracy (Deloire 2021). The same year, the Forum on Information and Democracy was formed. Established within its framework and led by Maria Ressa and Marietje Schaake, the Working Group on Infodemics offered several specific recommendations on the information and communication space (Report of Forum ID 2020).

The first group of recommendations addresses the need for public regulations governing the sector. For one thing, these would force Internet service providers to be transparent. First, transparency requirements should apply to all core digital platform functions within the public information ecosystem – content moderation, content ranking, content targeting and social influence building. Second, regulators responsible for enforcing transparency requirements should be able to exercise robust democratic control over these entities and have them audited. The third point is that sanctions for non-compliance should entail substantial fines, compulsory disclosures about the sanctions, possible legal consequences for CEOs, and administrative measures, such as denying access to a given country's market.

The second set of suggestions recognises the need for a new model of meta-regulations on content moderation. For platforms, this means the obligation to follow the rules enshrined in human rights – primarily equality and non-discrimination. Furthermore, they should fulfil the same pluralism requirements as radio and television broadcasters. Platforms must hire more moderators and allocate a certain percentage of their income to improve their content monitoring capabilities. The third group of recommendations calls for a new approach to designing social networking platforms. The central concept here is to establish a Digital Standards Enforcement Agency to enforce safety and quality standards of digital architecture and software engineering. The Forum on Information and Democracy has declared its readiness to commence work on a feasibility study for such an agency. Another proposition is that all conflicts of interests of platforms should be prohibited to avoid the information and communication space being governed or influenced by commercial, political or any other interests. Moreover, a co-regulatory framework – based on self-regulatory standards – should be established to promote journalism that serves the public interest.

The last group of recommendations calls for safeguards in closed messaging services when they enter a public space logic. These would limit some functions to curb the virality of misleading content by imposing opt-in features to receive group messages and measures to combat bulk messaging and automated behaviour. Furthermore, Internet

service providers should be more diligent in informing users of the origin of the messages they receive, especially those that have been forwarded. Finally, platforms should reinforce notification mechanisms of illegal content by users and appeal mechanisms for those users who were banned.

6 EU regulations

Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services and amending Directive 2000/31/EC (Digital Services Act 2020), passed by the European Parliament on 20 January 2022, is likely to play a critical role. The key significance of this piece of legislation primarily stems from its European-wide range (albeit "limited" to 27 EU Member States), and from its effectiveness and potential to set global trends. It appears that:

[...] a regulation instead of a directive is the right choice, better aligning with the goal of establishing harmonised and coherent regulations for the digital market across the EU. Previous experience implementing and applying the E-Commerce Directive in individual EU States and the growing adverse phenomena and activities online suggested the need to design laws more in tune with the ever-changing business and technological realities. This would serve to equalise competitive opportunities and give digital operators on the EU market a firmer legal ground, as well as preserve the "country of origin" principle. Consequently, there would be more incentive for European companies to develop their services and expand digitally in the UE beyond their domestic markets (PIIT 2021).

As noted in the rationale for the draft regulation, the reason for the changes is the conclusion that:

[…] since the adoption of Directive 2000/31/EC (the "e-Commerce Directive"), new and innovative information society (digital) services have emerged, changing the daily lives of Union citizens and shaping and transforming how they communicate, connect, consume and do business. […] At the same time, the use of those services has also become the source of new risks and challenges, both for society as a whole and for individuals using such services […] The coronavirus crisis has shown the importance of digital technologies in all aspects of modern life It has clearly shown the dependency of our economy and society on digital services and highlighted both the benefits and the risks stemming from the current framework for the functioning of digital services (Digital Services Act 2000).

Several important reasons have been stressed to explain the need for such regulations. First, they aim to provide effective solutions and mechanisms to counter illegal online content. Second, they are designed to create a fair and safe e-commerce environment. Third, their purpose is to ensure fairness in online advertising. Furthermore, it is necessary to regulate these three aspects at the EU – as opposed to national – level. It is also believed that the legislation would help protect the fundamental rights of citizens enshrined in the

EU Charter of Fundamental Rights. Measures will be proportionate to the type and size of the intermediary service providers and will include the gradation of their obligations. By placing systemic risk analysis obligations on online platforms, the regulation will facilitates risk management. It will also promote cross-border cooperation and, last but not least, afford a relatively firm legal ground to digital service providers, effectively driving the industry's growth (Soppa-Garstecka 2021: 5–9). In this context, it is stressed that, with these regulations in place, there will be greater democratic control and better monitoring of platforms, as well as a lower systemic risk of manipulation and disinformation.

Like Directive 2000/31/EC, the draft Regulation starts by defining instances under which Internet service providers are exempt from liability (Baran 2021: 19-27). This includes services such as mere conduit, caching and hosting. Such exemption from liability may also apply when voluntary proceedings are instigated on an own initiative basis. Similar to Article 15 of Directive 2000/31/EC, draft Article 7b stipulates that Member States may not impose on these providers a general monitoring obligation or an obligation to actively seek facts. However, they may establish obligations to counter illegal content (Article 8) and provide information (Article 9) to competent judicial and administrative public authorities. In Chapter III, the Regulation defines due diligence obligations for a transparent and safe online environment. Section 1 sets out obligations applicable to all providers of intermediary services: to designate a single point of contact to enable them to communicate directly, by electronic means, with Member States' authorities, the Commission and the Board (Article 10); providers which do not have an establishment in the Union but which offer services in the Union are obliged to designate a legal representative in the EU (Article 11); the obligation to include, in their terms and conditions, any restrictions that they impose regarding the use of their service in respect of information provided by the recipients of the service, and to act responsibly in terms of applying and enforcing these restrictions (Article 12); transparency reporting obligations for the removal of, or the disabling of access to, information considered illegal content or content incompatible with providers' terms and conditions (Article 13). Section 2 of this Chapter includes additional obligations applicable to hosting service providers. The plan is to make these providers obliged to introduce reporting mechanisms for alleged illegal content (Article 14); if a hosting service provider decides to remove certain information provided by the recipient of the service or disable access to it, it will be obliged to provide the recipient with a statement of reasons (Article 15).

The next part of the draft Regulation sets out further responsibilities. These, however, do not apply to micro-business or small-sized enterprise online platforms. All other platforms must ensure an internal complaint-handling system for decisions relating to alleged illegal content or information that is incompatible with their terms and conditions (Article 17). This includes the obligation of online platforms to cooperate with certified out-of-court dispute settlement bodies to resolve any conflicts with users of their services (Article 18). Furthermore, online platforms must prioritise notices submitted by trusted flaggers (Article 19) and take specific measures to counter inappropriate use (Article 20).

They are, additionally, required to inform law enforcement agencies of any suspicion of serious crimes involving a threat to the life or safety of persons (Article 21). Furthermore, online platforms are required to receive, keep, make the best efforts to assess the reliability of, and publish information about traders using their services, where such platforms allow consumers to conclude distance contracts with such traders (Article 22). Moreover, online platforms must organise their online interfaces so traders can comply with their obligations regarding pre-contractual information, compliance and product safety information under applicable Union law (Article 22(a)). They are also required to publish reports on their activities involving the removal of, and disabling access to, information considered illegal content or information that is incompatible with their terms and conditions (Article 23). This section also includes online platforms' obligations relating to online advertising transparency (Article 24). In the following part, the Regulation lays down obligations related to how so-called huge online platforms manage systemic risk (within the meaning of Article 25). They will be required to perform systemic risk assessments regarding the operation and use of their services (Article 26), take sound and effective measures to mitigate systemic risk (Article 27), and be subject to independent third-party audits (Article 28). Here, an additional obligation is imposed on very large online platforms using recommendation systems (Article 29) or having online advertisements displayed on their online interfaces (Article 30). What is more, the Regulation sets out the terms under which such content aggregators are to provide the Digital Services Coordinator of the establishment or the Commission, as well as vetted researchers, with access to data (Article 31). It also enforces the requirement to appoint compliance officers to ensure compliance with the obligations laid down in the Regulation (Article 32), in particular, additional transparency reporting obligations (Article 33).

Section 5 includes provisions on due diligence obligations – that is, processes in respect of which the Commission shall support and promote the development and implementation of harmonised European standards (Article 34); a framework for the development of codes of conduct (Article 35), and a framework for the development of detailed codes of conduct on online advertising (Article 36). The Regulation also contains a provision on crisis protocols for extraordinary circumstances which affect public safety and health (Article 37) (Soppa-Garstecka 2021: 27–29).

7 National regulatory attempts

We should mention that some countries have attempted to regulate liability for online content. For instance, in 1996 – before the emergence of social media – the US Congress passed the Communications Decency Act. Section 230 of this law protects online intermediaries against liability for user-published content, except for copyright infringements and child-trafficking content. Under Section 230, online platforms may also remove user speech. Also in the US, in September 2022, the State of California passed Assembly Bill 587, imposing specific transparency standards on social media platforms and making them subject to remedies in cases of disinformation, hate speech,

etc. (Assembly Bill (2022) No. 587). In September 2023, the X Corp (formerly Twitter) social media platform sued the State of California for passing this law, arguing that it represented a violation of the First Amendment of the Constitution of the United States, i.e., the right to freedom of speech and the Constitution of California (Case 2023).

Adopted in 2017, the German law *Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken* (*Netzwerkdurchsetzungsgesetz/Network Enforcement Act*, NetzDG, 2017) represents a particularly notable piece of legislation. Among other things, this Act introduced the obligation for social media networks to implement an effective and transparent complaints-handling procedure. In addition, these entities are required to publish transparency reports on implemented procedures, complaint figures and removed content. Moreover, the law obliges social networks with over two million registered users in Germany to remove "clearly illegal" content (such as posts, images, and videos) within 24 hours from notification. For content that is not obviously illegal, however, providers have a maximum of seven days to decide how to handle the case.

In Poland, on 15 January 2021, the Ministry of Justice put forward the Draft Act on the Protection of Freedom of Speech on Social Networking Services (Draft Acton the Protection of Freedom of Speech on Social Networking Services 2021). The legislation still awaits passage as the bill has not been submitted to the Polish Parliament. With this bill, the Polish legislators aim to create an environment that supports freedom of speech, ensures the right to receive true information, improves the protection of human rights and freedoms on social networking services that are accessible in the territory of the Republic of Poland and have at least a million registered users, as well as ensures that social media websites comply with the freedom to express opinions, acquire and disseminate information, express religious convictions, world views and philosophy of life, and the freedom to communicate. Under the draft Article 2, this normative act will: set out the rules for scrutinising businesses providing services by electronic means through social networking services with at least a million registered users such that it is possible for public authorities to guarantee that the users of these services enjoy their right to freedom of speech and access to factual information; define rules governing service providers' liability for publishing illegal content on social networking services, as well as service provider's obligations related to guaranteeing freedom of speech and access to factual information; specify rules under which service providers are to conduct internal control procedures to handle user complaints against content that is unlawful and against good morals, or that infringes the right to freedom of speech or access to true information; as well as rules governing proceedings before public administration bodies and court proceedings in the event of restricted access to an electronic service provided through a social networking service.

The statutory definitions of certain terms provided in draft Article 3 essentially determine the subjective and objective scope of this Regulation. Accordingly, the term "online networking service" is understood as a service provided by electronic means allowing users to share any content with other users or the general public and has at least a million

registered users in Poland. "Disinformation" is defined as false or misleading information produced, presented or disseminated for profit or against public interest. "Criminal content" means content that glorifies or incites certain prohibited acts (i.e., the acts referred to in Articles 117–119, 127–130, 133, 134–135, 137, 140, 148–150, 189–189a, 190a, 194–204, 222–224a, 249–251, 255–258, 343 of the Penal Code), or fulfils the definitional elements of a prohibited act. "Unlawful content" is content infringing personal interests, disinformation, criminal content, and content against good morals, including, in particular, content that promotes or glorifies violence, suffering or humiliation. Under the proposed law, "restricted access to content" means any acts and omissions to facilitate any forms of restricting access to content published in a social networking service, including removal of user-published content that is not unlawful and restricting access to content through algorithms or tags used by the service provider to indicate possible violations in the published content; "restricted access to the user profile" means removing or disabling access to the user profile, restricting or disabling the option to share content with other users on the user profile, including through the service provider's use of algorithms reducing the visibility of user-shared content or tags indicating possible violations in the published content.

Moreover, the bill envisaged the appointment of the Freedom of Speech Council – a public administration body watching over social networking services' compliance with laws governing the freedom to express opinions, to acquire and disseminate information, to express religious convictions, world views and philosophy of life, and also the freedom to communicate. The Council would serve in six-year terms, and its members would be allowed to stand for re-election to further terms. According to the proposal, the Council's chair would be elected by the Polish Parliament with a 3/5 majority of votes, subject to at least half of the statutory number of MPs being in attendance. If none of the candidates receives the 3/5 majority of votes, there is a revote, with the Polish Parliament appointing the chair with a simple majority. Council members would also be appointed by the Polish Parliament with a 3/5 majority vote, subject to at least half of the statutory number of MPs being in attendance; however, should a candidate for Council membership fail to receive the 3/5 majority of votes in the first vote, or if there is more than one candidate for Council membership and none of the candidates has received the 3/5 majority of votes, there would be a revote at the Polish Parliament, except that this time a simple majority would decide the result.

The proposed law lists several new obligations for social media platforms. First, the draft Article 15 prescribes that service providers receiving more than 100 user complaints – on account of their providing access to unlawful content, restricting access to content or restricting access to user profiles – per calendar year will be required to issue biannual reports, in Polish, to disclose how these complaints were dealt with. These reports would be published on the respective social networking services a month after the end of the relevant half-year at the latest. Second, service providers would have an obligation to designate one or more – but no more than three – national representatives to transact on their behalf all court and out-of-court business, handle complaints in internal control

procedures and provide institutions and authorities with any answers and information they may request for the purposes of their proceedings. Moreover, under the bill, service providers would have to implement effective and intelligible internal control procedures in Polish to handle matters raised in user complaints. Users dissatisfied with how their complaints were handled in internal control procedures would have the option to complain to the Council. After completing its complaint procedure, the Council would make a decision with which it would either order the provider to restore access to the restricted content or user profiles – on account of its finding that such restricted content or profiles do not represent unlawful content – or refuse to restore access to the restricted content or user profiles on account of its finding that they represent unlawful content.

Under Article 29 of the draft Act, service providers would not have the right to yet again restrict access to the content examined by the Council. According to the proposed law, service providers breaching the Act would face fines ranging from PLN 50,000 to PLN 50 million. Specifically, such fines would be imposed on service providers defaulting on their obligation to: 1) issue the report referred to in Article 15(1); 2) designate the national representative referred to in Article 16(1); 3) immediately notify the President of the Office of Electronic Communications about the designation or replacement of the national representative, stating their personal details as referred to in Article 16(3); 4) immediately notify the President of the Office of Electronic Communications about any changes in the personal details referred to in Article 16(4); 5) publish the complete personal details, as referred to in Article 16(5), in its social networking service in such a manner that they are clearly visible and directly and permanently accessible; 6) provide the individuals involved in internal controls with the training referred to in Article 17(1); 7) implement an effective and intelligible internal control procedure in Polish to handle the matters referred to in Article 19(1); 8) publish in its social networking service the Rules and Regulations of the service, accessible by all users and setting out the internal control procedure referred to in Article 19 (2); 9) ensure a clearly visible , directly and permanently accessible method of sending complaints in internal control procedures, as referred to in Article 19(3); 10) comply with the Council's decision ordering that the restricted access to content or user profile be restored, as referred to in Article 25(1); and 11) comply with the prosecutor's decision ordering that access to criminal content, as referred to in Article 37(2), be disabled. Service providers would also be subject to fines should their national representatives fail to comply with the obligations to 1) handle a user complaint in an internal control procedure, in the manner referred to in Article 20; 2) provide institutions and authorities with any answers and information they may request for their proceedings; and 3) participate in training courses organised by the President of the Office of Electronic Communications and concerning the current legal situation regarding user complaints handled in internal control procedures.

The party dissatisfied with how the matter was resolved may request reconsideration by the Council. If any criminal content is identified, the prosecutor may request the service provider or its national representative to provide any necessary information, including, in particular, user identification data and the relevant publications posted on the social

networking service. If the criminal content is found to include pornographic content involving minors or content glorifying or inciting terrorist acts, to entail the risk of serious harm or to cause difficult-to-remedy consequences when left accessible, the prosecutor may immediately order the service provider to disable access to such content. Additionally, the party concerned would have the option of complaining against the prosecutor's decision, with the district court having jurisdiction over the prosecutor's office which issued the contested decision. On a side note, the bill's final provisions propose several useful solutions. One of them would be particularly welcome – the John Doe lawsuit – a special lawsuit filed against an unidentified defendant to protect personal interests.

8 Concluding remarks

At this point, the necessity for certain regulations governing social media is absolutely clear. However, regulators must proceed with considerable caution regarding the scope, method and range of regulations. Proponents of regulating this area have expressed concern that regulators might feel tempted to restrict freedom of speech and legislate measures that effectively censor and stifle public debate. Still, if there were any doubts about the need for regulatory measures involving social networking platforms, there should be none by now. This includes international, regional and national regulations alike. Finally, there is the question of which type of regulations would be the most appropriate in this case. Would hard law or soft law in combination with self-regulatory measures work enough? Or would it be better to focus more on promoting social media safety skills? Perhaps the best solution would be to combine all three.

References:

Balkin, J.M. (2018) How to regulate (and not regulate) social media, *Journal of Free Speech Law*, available at: https://www.journaloffreespeechlaw.org/balkin.pdf (February 17, 2024).

Baran, A. (2021) Wyłączenie odpowiedzialności dostawców usług pośrednich w świetle Aktu o usługach cyfrowych, In: Konarski, X., Wasilewski, P. & Baran, K. (eds.) *Akt o usługach cyfrowych. Kompendium*, available at: https://www.piit.org.pl/__data/assets/pdf_file/0022/17707/akt_o_uslugach_cyfrowych_kompendium.pdf (February, 17 2024).

Barrett, P. (2020) *Regulating Social Media*, available at: https://static1.squarespace.com/static/5b6df958f8370af3217d4178/t/5f58df637cbf80185f372776/1599659876276/NYU+Section+230_FINAL+ONLINE+UPDATED_Sept+8.pdf (February 20, 2022).

Brannon, V.C. (2019) *Free Speech and the Regulation of Social Media Content*, available at: https://sgp.fas.org/crs/misc/R45650.pdf (February 20, 2022).

Briggs, S. (2018) *The Freedom of Tweets. The Intersection of Government Use of Social Media and Public Forum Doctrine*, available at: http://blogs2.law.columbia.edu/jlsp/wp-content/uploads/sites/8/2019/04/Vol52-Briggs.pdf (February 20, 2022).

Case (2023) *X CORP. v. Bonta District Court, E.D. California*, available at: https://www.courtlistener.com/docket/67776396/x-corp-v-bonta/ (February 17, 2024).

Diaz, J. (2021) *Jack Dorsey Says Trump's Twitter Ban Was 'Right Decision' But Worries About Precedent*, available at: https://www.npr.org/2021/01/14/956664893/twitter-ceo-tweets-about-banning-trump-from-site?t=1645378363608 (February 17, 2022).

Digital (2022) *Global Overview Report*, available at: https://datareportal.com/reports/digital-2022-global-overview-report (Febaury 25, 2022).

Dudek, W. (1991) *Międzynarodowe aspekty mass mediów* (Katowice: Uniwersytet Śląski).

Dziemba, R. (2009) Nowe media a koncepcja „piątej władzy", In: Jeziński, M. (ed.) *Nowe media i polityka. Internet, demokracja, kampanie wyborcze* (Toruń: Wydawnictwo Adam Marszałek), pp. 53-61.

Fagan, F. (2018) Systemic Social Media Regulation, *Duke Law & Technology Review*, available at: https://scholarship.law.duke.edu/dltr/vol16/iss1/14/ (Febaury 17, 2022).

Fox, Ch. (2021) *Social Media. How Might It Be Regulated?*, available at: https://www.bbc.com/news/technology-54901083 (Febaury 20, 2022).

Garcia, R. & Hoffmeister, Th. (2017) Social Media Law in a Nutshell, *School of Law Faculty Publications*, (21), available at: https://ecommons.udayton.edu/law_fac_pub/2 (Febaury 17, 2024).

Gardocki, L. (1993) Europejskie standardy wolności wypowiedzi a polskie praw karne, "*Państwo i Prawo*", (3), pp. 111–156.

Garlicki, L. (1997) Art. 10. Wolność wyrażania opinii, In: Gardocki, L. (ed.) *Konwencja o Ochronie Praw Człowieka i Podstawowych Wolności. Komentarzt. 1*. (Warszawa: Wydawnictwo C.H. Beck), pp. 583-648.

Grzesik-Robak, A. (2009) Media tradycyjne wobec nowych rozwiązań technologicznych – szansa czy zagrożenie, In: Jeziński, M. (ed.) *Nowe media i polityka* (Toruń: Wydawnictwo Adam Marszałek), pp. 160-165.

Hudson, D.L.Jr. (2017) *Hate Speech Online*, available at: https://www.freedomforuminstitute.org/first-amendment-center/primers/fake-news-primer/ (Febaury 17, 2024).

Jabłoński, M. & Wygoda, K. (2002) *Dostęp do informacji i jego granice* (Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego).

Jaskiernia, A. (2005) Swoboda debaty politycznej w mediach w świetle standardów Rady Europy, *Studia Medioznawcze*, 4(23), pp. 90-104.

Kamiński, I.C. (2005) Media w europejskiej konwencji o ochronie praw człowieka i podstawowych wolności, In: Barta, J., Markiewicz, R. & Matlak, A. (eds.) *Prawo mediów* (Warszawa: Wydawnictwo LexisNexis), pp. 33-75.

Kamiński, I.C. (2009) Karta Praw Podstawowych jako połączenie praw i zasad – strukturalna wada czy szansa?, In: Wróbel, A. (ed.) *Karta Praw Podstawowych w europejskim i krajowym porządku prawnym* (Warszawa: Wydawnictwo Wolters Kluwer Polska), pp. 38-39.

Kaplan, A.M. & Haenlein, M. (2010) Users of the World, Unite! The Challenges and Opportunities of Social Media, *Business Horozins*, 53(1), pp. 59-68, https://doi.org/10.1016/j.bushor.2009.09.003.

Kayode-Adedeji, T., Oyero, O. & amp; Aririguzoh, S. (2017) Regulating the Social Media for Global Relationships, *"New Trends and Issues Proceedings on Humanities and Social Sciences"*, 4(10), pp. 426-433.

Kaznowski, D. (2013) Social media – społeczny wymiar Internetu, In: Królewski, J. & Sala, P.(eds.) *Marketing. Współczesne trendy. Pakietstartowy* (Warszawa: Wydawnictwo PWN), pp. 81-103.

Khan, I. (2021) *How Can States Effectively Regulate Social Media Platforms?*, available at: https://www.law.ox.ac.uk/business-law-blog/blog/2021/01/how-can-states-effectively-regulate-social-media-platforms (Febaury 17, 2024).

Kowalski, T. (2003) Wprowadzenie do analizy struktury przemysłu nowych mediów, *Studia*

Medioznawcze, 3(13), pp. 23-30.
Kreft, J. (2013) *Władza platform. Za fasadą Google, Facebooka i Spotify* (Kraków: Universitas).
Kreft, J. (2017) *Koniec dziennikarstwa jakie znamy. Agregacja w mediach* (Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego).
Krzysztofek, K. (2007) Status mediów cyfrowych: stare i nowe paradygmaty, In: Fras, J. (ed.) *Studia nad komunikacją popularną, międzykulturową, sieciową i edukacyjną* (Toruń: Wydawnictwo Adam Marszałek), pp. 83-99.
Kumm, M. (2023) *Market Imperatives and the Public Sphere: Constitutional issues concerning the regulation of social media*, available at: https://www.wzb.eu/en/events/market-imperatives-and-the-public-sphere-constitutional-issues-concerning-the-regulation-of-social (February 17, 2024).
Mikułowski-Pomorski, J. (1988) *Informacja i komunikacja. Pojęcia, wzajemnerelacje* (Ossolineum: Wrocław-Warszawa-Kraków-Gdańsk-Łódź).
Młynarska-Sobaczewska, A. (2003) *Wolność informacji w prasie* (Toruń: Towarzystwo Naukowe Organizacji i Kierownictwa Dom Organizatora).
Młynarska-Sobaczewska, A. & Preisner, A. (2008) Rozwój technologiczny a przyszłość demokracji, In: Gizicka, D. & Gizicki, W. (eds.) *Człowiek i społeczeństwo wobec wyzwań współczesności. Aspekty kulturowe i społeczne* (Toruń: Wydawnictwo Adam Marszałek), pp. 34-45.
NetzDG (2017) *Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken* (*Netzwerkdurchsetzungsgesetz*, available at: https://www.gesetze-im-internet.de/netzdg/BJNR335210017.html (February 23, 2022).
Ohlheiser, A. & Guo, E. (2021) *Twitter Locked Trump's Account. Insiders Say It Needs to Go Further*, available at: https://www.technologyreview.com/2021/01/06/1015830/twitter-trump-suspension-ban/ (February 17, 2024).
Oleksiuk, I. (2000) Wolność wypowiedzi w Internecie, *Studia Medioznawcze*, (1), pp. 97-108.
Palmer, A. (2021) *Facebook Will Block Trump from Posting at Least for the Remainder of His Term*, available at: https://www.cnbc.com/2021/01/07/facebook-will-block-trump-from-posting-for-the-remainder-of-his-term.html (February 17, 2024).
Palmer, J. W. & Eriksen, L. (1999) Digital News – Paper, Broadcast and More Converge on the Internet, *The International Journal on Media Management*, 1(1), pp. 125-135.
Papiernik, O. (2021) Sygnalistka nie zostawia suchej nitki na Facebooku. Zagraża zdrowiu, *Dziennik.pl.*, available at: https://technologia.dziennik.pl/sprzet/artykuly/8287958,sygnalistka-facebooku-zagrozenie-zdrowie-zarzuty.html (February 23, 2022).
Paslawsky, A. (2017) The Growth of Social Media Norms and the Governments' Attempts at Regulation, *Fordham International Law Journal*, 35(5), pp. 1486-1490.
Patterson, K. (2020) *Why regulating the public sphere matters more than ever*, available at: https://pec.ac.uk/blog/why-regulating-the-public-sphere-matters-more-than-ever (February 17, 2024).
PIIT (2021) *Stanowisko Polskiej Izby Informatyki i Telekomunikacji w sprawie projektu regulacji Digital Services Act – dokument COM(2020) 825*, available at: https://www.piit.org.pl/__data/assets/pdf_file/0021/17238/PIIT_-Stanowisko_regulacje-DSA_22.01.2021.pdf (February 17, 2024).
Raport of Forum ID 2020 (2020) Forum of Information & Democracy, *Working group of infodemics*, available at: https://informationdemocracy.org/wp-content/uploads/2020/11/ForumID_Report-on-infodemics_101120.pdf (February 17, 2024).
Report (2021) *Online hate increasing against minorities*, available at: https://www.ohchr.org/en/stories/2021/03/report-online-hate-increasing-against-minorities-says-expert (February 17, 2024).

Ressa, M. (2020) *Report on Infodemics. Working Group on Infodemics. Policy Framework*, available at: https://informationdemocracy.org/wp-content/uploads/2020/11/ForumID_Report-on-infodemics_101120.pdf (February 17, 2024).

Sadurski, W. (1998) Wolność prasy w systemie praw człowieka (wybrane zagadnienia), In: Zubik, M. (ed.) *Obywatel – jego wolności i prawa. Zbiór studiów z okazji 10-lecia Rzecznika Praw Obywatelskich* (Warszawa: Wydawca Biuro RPO), pp. 75-93.

Sadurski, W. (1992) Prawo do wolności słowa w państwie demokratycznym, *Państwo i Prawo*, (10), pp. 3-10.

Scaife, L. (2018) *The Effective Regulation of Social Media*, available at: http://kar.kent.ac.uk/65920 (February 17, 2024).

Skrzypczak, J. (2019) The Right to Freedom of Opinion and Expression in the Universal Declaration of Human Rights – a Contemporary Perspective, In: Sungurov, A., Fernández Liesa, C.R., Barranco Avilés, M. del C., LlamazaresCalazadilla, M.C. & Pérez De La Fuente, Ó (eds.) *Current Issues on Human Rights* (Madrid: Dykinson), pp. 81-92.

Sobczak, J. (2008) *Prawoprasowe. Komentarz* (Warszawa: Wolters Kluwer).

Soppa-Garstecka, M. (2021) 10 powodów, dla których Europa potrzebuje Aktu o usługach cyfrowych, In: Konarski, X., Wasilewski, P. & Baran, K. (eds.) *Akt o usługach cyfrowych. Kompendium*, available at: https://www.piit.org.pl/__data/assets/pdf_file/0022/17707/akt_o_uslugach_cyfrowych_kompendium.pdf (February 17, 2024).

Soppa-Garstecka, M. (2021) Nakaz podjęcia działań przeciwko nielegalnym treściom oraz nakaz udzielania informacji, In: Konarski, X., Wasilewski, P. & Baran, K. (eds.) *Akt o usługach cyfrowych. Kompendium*, available at: https://www.piit.org.pl/__data/assets/pdf_file/0022/17707/akt_o_uslugach_cyfrowych_kompendium.pdf (February 17, 2024).

Stahl, T. (2020) Privacy in Public: A Democratic Defense, *Moral Philosophy and Politics*, https://doi.org/10.1515/mopp-2019-0031.

Tan, C. (2018) *Regulating Content on Social Media: Copyright, Terms of Service and Technological Features* (UCL Press), https://doi.org/10.2307/j.ctt2250v4k.

Toffler, A. (2002) Odmasowione środki przekazu. Maryla Hopfinger (przeł.), In: Hopfinger, M. (ed.) *Nowe media w komunikacji społecznej w XX wieku* (Warszawa: Wydawnictwo Oficyna Naukowa), pp. 441-447.

Tully, S. (2014) People You Might Know. Social Media in the Conflict Between Law and Democracy, In: Patmore, G. & Rubenstein K. (eds.) *Law and Democracy. Contemporary Questions* (Canberra: ANU Press), pp. 53-172.

Veltman, K.H. (2006) *Understanding New Media* (Calgary: Augmented Knowledge & Culture).

Wall, J., Mooppan, H., Joshi, S., Taibleson, R., Mc Intosh, S. & Utrecht, J. (2020) *Donald J. Trump, v. Knight First Amendment Institute At Columbia University*, available at: https://www.supremecourt.gov/DocketPDF/20/20-197/150726/20200820102824291_Knight%20First%20Amendment%20Inst.pdf (February 17, 2024).

Wasikowski, T. (1983) Informacja i komunikowanie masowe w stosunkach międzynarodowych, *Sprawy Międzynarodowe*, (2), pp. 75-93.

Zafarani, R., Abbas, M. & Liu, H. (2014) *Social Media Mining. An Introduction* (Cambridge: University Press).

Chapter IV

Information, Disinformation, Cybersecurity

KATARZYNA CHAŁUBIŃSKA-JENTKIEWICZ & MONIKA NOWIKOWSKA

Abstract Information has been and is an integral part of all human activity. For any state, it is a strategic commodity, where every plane of security depends on the information resource it possesses, which must be protected accordingly. It can be hypothesised a priori that the protection of important information resources is inextricably linked to the security interests of the state and its subjects. Without doubt, the most sought-after resource today is information. The introduction of information and computer technologies and their increasingly widespread use has led to a situation where there is an unfettered exchange of information between remote entities and logistical considerations do not matter. The turbulent and very dynamic development of information technology, as well as the rapidly increasing amount of data being processed, has necessitated the search for solutions to effectively manage information, taking into account the risks involved, especially in cyberspace.

Our technological capabilities are steadily advancing, but this is not always a reason to rejoice. It is important to be aware not only of the benefits of this, but also of the risks. In this case, we can speak of new phenomena such as disinformation, deepfake, fake news and trolling. Until recently, the pinnacle of disinformation was the dissemination of fake photos and texts. However, with the development of the digital age, the possibilities of artificial intelligence have also developed, which has reached a whole new level and is now also able to create fake videos. With the development of deepfake technology, a breeding ground has emerged for the spread of disinformation in the political sphere and for influencing public opinion regarding specific public office holders. Fake information can now

CORRESPONDENCE ADDRESS: Katarzyna Chałubińska-Jentkiewicz, Ph.D., Associate Professor, Kozminski University, College of Law, ul. Jagiellonska 57/59, 03-301 Warszawa, Poland, e-mail: kasiachalubinska@gmail.com, ORCID: 0000-0003-0188-5704. Monika Nowikowska, Ph.D., Assistant Professor, War Studies University in Warsaw, Faculty of Law and Administration, Al. Generała Antoniego Chruściela „Montera" 103, 00-910 Warszawa, Poland, e-mail: m.nowikowska@akademia.mil.pl, ORCID: 0000-0001-5166-8375.

https://doi.org/10.4335/2024.2.4 ISBN 978-961-7124-25-5 (PDF)
Available online at http://www.lex-localis.press.

be used to influence the electoral process, poor media reception, social tensions, acts of unfair competition and also other disinformation attacks that disrupt the normal functioning of the state and individuals. This now creates a whole new level of risks associated with the spread of false information.

Keywords: • cybersecurity • disinformation • fake news • information • information society • information warfare

1 General comments

The use of modern techniques has led to the emergence of a new type of society, which is called information society (Chałubińska-Jentkiewicz, Nowikowska 2020:34). Information society is a society that has the technical and legal instruments and the knowledge to use these instruments (Chałubińska-Jentkiewicz, Karpiuk, 2015:39). It is a society for which information is the most important value and commodity, which is sought, *inter alia*, through the Internet. Thanks to the Internet, people have gained the possibility to access unlimited sources of information. However, with the rapid spread of information, the possibility of manipulating messages or creating false ones has also emerged. A phenomenon often called "information warfare" has also emerged (Wasiuta, Wasiuta 2017: 71), referring to influencing the civilian and/or military population of another country by disseminating appropriately selected information. The object of information warfare is both collective and individual awareness while information influence can both against a background of information noise and in an information vacuum (Nowikowska, 2023: 64). The introduction of foreign objectives makes information warfare a war and distinguishes it from mere propaganda. The resources of information warfare are various communication tools – from the media to email and gossip. Information includes distorting facts or imposing on citizens an emotional perception that is convenient for the aggressor (Formicki, 2017: 319).

The definition of information refers to the concept of data (spatiotemporal events, states, procedures, numbers and descriptions) that relate to a model (a system, a slice of the world with real and unreal components, relationships) and, at the same time, perpetuate (describe, constitute, change) this model from the point of view of a person (data user) and for a specific purpose. Data become information only by relating them to exact situations in the model (reality) from the user's perspective and for a purpose. Information (Latin *informatio* – presentation, image; *informare* – to shape, present) is an interdisciplinary term, defined differently in various fields of science, the essence of which is the reduction of uncertainty (indeterminacy) (Chałubińska-Jentkiewicz, Nowikowska 2020: 22).

Key concepts in the communication process include manipulation (Latin *manipulatio* – manoeuvre, foray, trick; *manus* – arm, *manipulus* – hand), which is a form of influencing a person or group in such a way that they unconsciously and voluntarily pursue the goals of the manipulator. The ability to boss others around, the knowledge of how to be in charge, and how to conduct negotiations to get a partner to change their mind. It is often an inspired social interaction aimed at getting a person or a group of people to act contrary to their well-understood interests. Usually, the person or the group of people subjected to manipulation are unaware of how they are influenced. The author of the manipulation usually seeks personal, economic or political gain at the expense of those subjected to manipulation.

One can manipulate content (in the case of disinformation, it would be fake news) or how information is conveyed (in the case of disinformation, it would be manipulating the message so that genuine information is perceived falsely by building false opinions and positions, and drawing false conclusions). Linguistic manipulation is often used in propaganda. Although manipulation is perceived as unethical, it is often used in business relationships and negotiations. Manipulation is a modern technique of digital marketing communication.

By contrast, persuasion (Latin *persuasio*) is the skill of persuading someone that you are right about something. It differs from manipulation in that persuading someone to do something will not harm them, but the methods that are used when persuading are also employed when manipulating. Persuasion is also described as one of the methods of rhetoric or just as a reference to the "rhetorical tradition". It appeals to one's intellect, emotions and will. Considering the aims of persuasion, we can distinguish between:

- persuasion aimed to convince – it is to prove that something is right or true; it is the "purest" type of persuasion; it assumes that the recipient is a communicatively active individual and that the sender's intentions are honest and reliable;
- persuasion aimed to induce (propaganda) – it is to get as many supporters as possible for an idea or doctrine; it is a conscious effort to influence the recipient;
- persuasion aimed to incite (agitation) – it is to win the recipient over to an idea, cause or view.

Manipulating polls, results, or opinions has become the standard. The editors of the programme *Strefa starcia* (the Clash Zone) asked their viewers on Twitter whether they accepted the possibility for homosexual couples to adopt children. The results were to be presented on TVP Info, and the poll's authors probably hoped that the viewers' answers would give them a strong argument to fight against LGBT communities. Unfortunately for the poll's originators, the result fell short of expectations, so instead of revealing the viewers' vote, they decided to remove the question from the network. Instead, the following laconic message appeared on Strefa starcia's Twitter account: "We respect the votes of VIEWERS. Therefore, it is essential to us that the result of the poll reflects THEIR point of view and not that of bought-off farms of anonymous trolls".

A team of experts from EU Member States developed the official EU definition of disinformation, according to which this term "includes all forms of false, inaccurate, or misleading information designed, presented, and promoted to intentionally cause public harm or for profit" (https://www.cyberdefence24.pl/ue-unijna-definicja-dezinformacji-i-nowy-kodeks-postepowania-dla-mediow [accessed on: 02/06/2022]).

According to the above definition, disinformation is a deliberate action to evoke a specific social, economic or political reaction. Disinformation undermines trust in public institutions and harms democracies by making it difficult for citizens to make informed decisions. False information sows uncertainty and contributes to social tensions, having

potentially serious implications, particularly for public security and order. The development of modern technologies has made it easy for such false information to spread globally using any of the techniques indicated above to influence the recipient of the information (KRRiT 2020: 7).

It should be emphasised that disinformation is a situation in which information, although true and adequately communicated to the public, is intended to elicit false opinions and conclusions. Fabricating such a message by creating various types of false documents, organisations, etc., is misleading (produces an image of the world that is inconsistent with reality) and produces certain effects such as making wrong decisions by the recipient, forming a view, action or inaction, according to the assumption of the disinformer.

According to disinformation model analyses, from the perspective of cognitive models, disinformation is the effect of the influence of an imposed cognitive environment (worldview). Therefore, it can also be produced using a message that is highly congruent with the facts and a message that is consistent with the facts but evokes a false opinion or position about them.

Another important phenomenon, in the context of the issues discussed, is the publication of false images and sounds in the media. Unreal, modified photographs, so-called fake photos, which convey a false story, do not surprise the public any more. So-called fake videos have also gained popularity. Researchers have developed software that makes it possible to reconstruct the facial expressions of any person and create an image that matches lip movement to any text so that video content can also carry a false message. One of the companies dealing with this is Storyful, whose activities include verifying the multimedia posted by users on social networking sites, which are then used by the media worldwide (Reconstructing facial movement in real-time with a webcam? Nothing simpler, all you need is a webcam, https://whatnext.pl/ rekonstrukcja-ruchu-twarzy-czasie-rzeczywistym-przy-pomocy-kamery-internetowej-nic-prostszego/ [accessed on: 02/06/2022]).

The general understanding of the term *disinformation* differs from that contained in the literature. Disinformation affects entire populations rather than individuals. The development of disinformation is linked to the development of social communication techniques, and thus the freedom that corresponds to access to information and the right to disseminate it. Disinformation is, therefore, a side-effect of the colonialism of the web.

2 Post-truth era in the media

A new phenomenon changing existing communication process rules is the so-called post-truth. One of the most significant global crises of our time, involving the spheres of political, social, and cultural relations and, later, the scope of mass communication, has been named post-truth. In 2016, the editors of the Oxford Dictionary declared post-truth

the "word of the year". Such interest in the neologism is understandable, given the phenomenon which this word denotes. According to the Oxford Dictionary, post-truth is defined as an adjective "relating to circumstances in which objective facts are less influential in shaping public opinion than emotional appeals". The neologism became particularly popular in Western (primarily American and British) journalism when, back in 2016, Donald Trump won the US presidential election, and the British people voted for an exit from the European Union. Leading columnists of the major mass media of both countries heralded the beginning of the "post-truth era". Awareness of the new problem plunged the Western media into profound pessimism. In an information world, where actual data cannot be relied upon, and where evidence and testimony are no longer the main means of social dialogue, journalism is losing its relevance and, with it, the entire hitherto traditional value system of the media.

"Facts held a sacred place in Western liberal economies. Whenever democracy seemed to be going awry, when voters were manipulated, or politicians were ducking questions, we turned to facts", wrote British political economy professor W. Davis (Chałubińska-Jentkiewicz, 2023: 265). Although the word *post-truth* has become topical in the media discourse and has found its way into everyday language, it originated much earlier. It was coined by the American playwright of Serbian origin, S. Tesich, who, back in 1992, in an essay published in *The Nation* magazine, wrote about the political atmosphere in the USA: "...we (meaning Americans), as free people, have freely decided that we want to live in some post-truth world" (Tesich, 1992: 12). The word came into academic circulation later, in 2004, with the study titled "The Post-Truth Era: Dishonesty and Deception in Contemporary Life" (Keyes, 2022: 23). The author of the work, American researcher and writer Ralph Keyes, made a significant attempt to explore the reasons that have led modern society to a situation where truth has lost its fundamental meaning. "When our behaviour conflicts with our values, what we're most likely to do is reconceive our values" (Keyes, 2022: 25). It is clear that R. Keyes considers post-truth in a much broader sense than the political sphere and mass communication; in his case, it begins with interpersonal relationships. He points out that contemporary man has an "alternative ethic" that allows him not to suffer psychological distress when he lies. To justify oneself in the modern language, there are "transitional" phases between truth and falsehood, i.e., "alternative truth", "my truth", "this is how I see it", and "an alternative version of reality" (Chałubińska-Jentkiewicz, 2023: 266). R. Keyes draws attention to the processes that have led the world to post-truth. The author points to the influence of the philosophy and aesthetics of postmodernism, which has spread in mass culture through works of art, literature, cinema, etc. Postmodernism, as it is known, basically contains relativism, indifference to the problem of separating truth from falsehood or even insists on the impossibility of such a separation. Another source of influence specific to US culture is the so-called new journalism that emerged in 1970 (Weingarten 2017: 20). The development of "new journalism" not only helped to enrich the texts of newspapers and magazines with novelistic techniques but also to directly incorporate non-existent, made-up events and situations into journalistic texts. In fact, reporters began to use conscious

falsification of facts "to make stories beautiful". Obviously, this approach gradually erodes the reader's trust in the journalistic text. One more driver that leads Western culture towards post-truth is the film industry. R. Keyes points to the distortion of facts in films about real events and the mythologisation of the lives of film actors (Keyes, 2022: 25).

The problem of the "crisis of fact" is that, in the 21st century, we are faced with an overabundance of material that contains facts. Until then, fact was the basis for an objective description of the world. According to cultural historian Mary Poovey, this belief began to take shape in the Middle Ages when accounting emerged among traders, contributing to the development of science in subsequent years (Poovey 1998: 35). The 20th century introduced a new discipline – the use of numbers, which marketing companies and politicians quickly adopted. Today, an enormous number of sources, various forms of information transmission, and the inability to check the veracity of messages have resulted in a lack of confidence in the material that contains facts.

From a society based on facts, civilisation is entering the era of a society based on data. These are collected automatically through various devices and applications that determine user behaviour. The function of such data (big data) differs dramatically from the classic function of facts. If a fact served as proof in the public dialogue, in searching for an optimal solution, the data showed the public mood, making it possible to predict its behaviour and to adapt to its tastes and expectations. The fact as a testament to reality loses its value for the communicator: why prove something if you can count on recipients' preferences and offer them a narrative that will be accepted with great confidence? Textual modes of communication further facilitate the process of lying. Thus, a key factor in the spread of post-truth is technology, technical devices and their development (Chałubińska-Jentkiewicz 2023: 267).

Some researchers have highlighted the impact of biased online websites. Such media projects are not oriented towards adhering to the professional standards of journalism (balancing opinions, verifying facts, separating them from beliefs, etc.) (Nowikowska, 2020: 132). They have never been integrated into the system of professional journalism, ethical standards are unknown to them, and they aim to stay in power at all costs. The audience of any such site may be relatively small, but the information they produce can spread extremely quickly and widely through social networks. In addition, the actions of single sites with small audiences can have a cumulative effect and, as a result of the impact, have fairly significant audience support. It is no coincidence that social media has become a provider of disinformation and a "battlefield" for post-truth.

Firstly, the pattern of distribution of information – through subscriptions by "friends" – lowers the level of critical perception of news. The user sometimes does not even read the news to the end and presses the "share" button simply because of a pre-formed liking for the source of the information. Secondly, the administration of social networks itself is

not motivated to tackle such phenomena. As a result of the combined effect of old and new media, "fake" factories and social networks, a kind of "disinformation ecosystem" is created in the information space, in which the average reader cannot distinguish truth from falsehood. But it is vital to realise that, in the "disinformation ecosystem", the public is not a passive object of influence. Through emotions and manipulation by politicians and journalists, it is a participant in this system's processes, too. As a result of the spread of post-truth, the journalistic community has encountered the following problems: (a) the division of society based on the Self vs. Other principle, (b) the use of propaganda, manipulative techniques, emotional influence instead of a rationally balanced approach, (c) the emotional enlightenment of events, situations and problems, (d) the decline of the importance of information that contains facts, (e) the decline of the prestige of the media, journalists, experts and political activists, (f) the total distrust and at the same time the uncritical perception of information from sources that are recognised as the Self, and (g) the impossibility to have a full and constructive social dialogue (Chałubińska-Jentkiewicz, 2023: 268–269).

It should be noted that R. Keyes has warned of the danger of the emergence of a "suspicious society" in which the mechanisms of self-confidence will be destroyed – the more a person deceives himself, the more suspicious he becomes of others. This is a process of destroying public debate and democracy. However, R. Keyes also speaks of an existing "desire for righteousness", which can provide the basis for countering post-truth (Keyes 2022: 55). A situation of widespread threat that affects society as a whole and requires extraordinary countermeasures and focuses much of the attention and engagement of those in power is a circumstance conducive to disinformation. The problem with contemporary digital media is that we do not know how to define contemporaneity and what contemporary social media is anyway.

3 Fake news

Fake news is untrue or partly untrue information published, for example, on information services or social networking sites. Fake news aims to convince the recipient that it is information which describes the truth. However, it is important to distinguish false and misleading news from parody or satire, which is not intended to mislead the recipient. Fake news is also defined as false information, often of a sensationalist nature, published in the media to mislead the recipient for financial, political or prestige gain. Fake news can be an element of disinformation as part of measures described as active measures in the "black" hybrid technology group (Chałubińska-Jentkiewicz 2023: 269).

According to media expert Martina Chapman, all fake news consists of three elements or, in other words, its author intends to induce in its recipients a state of (1) suspicion, (2) so that they can be misled, and (3) manipulated (https://www.webwise.ie/teachers/what-is-fake-news [accessed on: 13/02/2024]).The following types of fake news can be distinguished:

1) clickbait – click lurkers and click baiters – websites that, using eye-catching thumbnails or sensational headlines, tempt people to visit them, taking advantage of people's natural curiosity and thus generating traffic to the website, increasing the number of clicks and revenue from the advertisements displayed on the pages;
2) biased/slanted news – on sites that present untrue news and which are visited by viewers seeking confirmation and reinforcement of their views, including prejudices;
3) satire and parody – found on sites that present untrue or exaggerated information for purely entertainment purposes;
4) sloppy journalism – information published by journalists without corroboration and without checking the credibility of sources (Chałubińska-Jentkiewicz 2023: 270).

The famous 18th-century Irish writer, essayist and satirist Jonathan Swift was to say: "Falsehood flies, and truth comes limping after it" (https://www.goodreads.com [accessed on 13/02/2024]. This well-known maxim from nearly three hundred years ago describes perfectly the effectiveness, speed and agility with which false information spreads in today's social media. At least this is the finding of the study titled "The spread of true and false news online" (Vosoughi, Roy, Aral 2018), conducted by researchers at the Massachusetts Institute of Technology, which was published in the *Science* magazine. In this study, researchers analysed the diffusion of true and false information spread on Twitter between 2006 and 2017. The dataset included nearly 126,000 pieces of information shared by 3–4.5 million site users. It turned out that fake news spread much faster, and reached further, deeper and more widely than true news, especially concerning information in the political sphere. It was also noted that the algorithms contributed equally to the diffusion of both factual and fake news, underpinning the conjecture that people, as social media users, are more likely to share negative information than positive information. The researchers also found that 1% of the most popular fake news reached between 1,000 and 100,000 recipients, with the figure rarely exceeding 1,000 recipients for 1% of the most popular true news. It is important to distinguish false and misleading fake news from satire or parody, which have a humorous function and are not intended to mislead the recipients (Nowikowska, 2020:88).

True news is also sometimes considered fake by individuals or institutions because of the negative content it carries for them. Recently, the use of this word has increased by 400%.

Fake news can be used in politics, e.g., during election campaigns. It is mainly aimed at eliminating the enemy and doing harm or discrediting the other side. One example of this is the US election. Between 2009 and 2013, Hilary Clinton was the Democratic Party's presidential candidate, and it was during this period that a vast amount of false information about her was created. One was Hilary Clinton's sale of weapons to terrorists (Palczewski, 2019: 144). This false information became the most popular at the time. It concerned the arms trade with ISIS, and the number of hits oscillated around 800,000 hits on Facebook.

Recently, NATO has also seen a huge influx of fake news and propaganda from Russia, especially after the annexation of Crimea. Fake news was created by Russian officials and distributed globally through media agencies. It could be located in popular and prestigious news services in the United States. In response, NATO developed an action strategy in 2019. The fight against disinformation continues to be an essential part of NATO's communication strategies and day-to-day operations, including media monitoring, information space analysis and proactive communication in a coordinated and fact-based manner. It aims to inoculate the media sphere instead of debunking every piece of false information (Chałubińska-Jentkiewicz 2023: 274).

"Response to Disinformation on COVID-19" was implemented in an action plan issued to allies by the Secretary-General. This document aimed to bring together multiple threads of work on countering hostile disinformation around COVID-19. In 2021, NATO's Toolbox for Countering Hostile Information Activities was created. It reflects a two-pronged response model through "understanding" and "engagement", supported by "coordination". NATO should also strengthen the mandate of existing bodies focusing on strategic communications to better coordinate national efforts. This could include sharing information on threats, incidents and best response practices, among other things. In 2021, the 2022 NATO Communications Strategy was produced. This document was presented to the North Atlantic Council on 14 December 2021. In the face of the war in Ukraine, the new 2022 Strategic Concept was adopted. At the summit held on 29–30 June 2022 in Madrid, NATO defined new objectives and directions for action, which also relate to activities in cyberspace and disinformation. The strategy envisages, among other things, digital transformation, adapting NATO's command structure to the information age and strengthening cyber defence, network and infrastructure. The strategy highlights that authoritarian actors challenge interests, values and the democratic way of life through disinformation campaigns.

In Poland, there is, *inter alia*, ISSA Poland – the Information Systems Security Association, which teaches network users how to counter disinformation (Jak-rozpoznawac-i-weryfikowac-faszywe-informacje-fake-news.png, https://pl.wikipedia.org/wiki/Plik:Jak-rozpoznawac-i-weryfikowac-faszywe-informacje- -fake-news.png [accessed on: 25/05/2022]).

In 2018, Kantar Public conducted studies on fake news. The study participants mostly indicated that they encountered false information shown in the media once a day, others once a week, a few times a month, rarely or never. Many people claimed to have the ability to recognise fake news and to distinguish it from reliable news (Fake news in Poland and Europe. The Kantar Public study, https://reporterzy. info/3634,fake-news-w-polsce-i-w-europie-badanie-kantar-public.html [accessed on: 25/05/2022]).

4 Trolling

Disinformation on the Internet can take the form of propaganda. This is a phenomenon commonly known as trolling. Internet trolls are a tool of disinformation. It is assumed that the Internet is an unbeatable power and information uploaded to the web will circulate. A troll is a person who knowingly posts thoughtful, mocking or provocative posts and comments on online forums. This action is intended to provoke discussion with other users. The main tasks of trolls include propaganda, manipulation, misrepresentation of facts, and introduction into popular awareness of short graphic information, so-called memes, which remain long in recipients' memories. Often, trolls act for money (they are paid to engage in this type of activity). Understood this way, trolling is an anti-social behaviour in the digital world (Chałubińska-Jentkiewicz, Nowikowska, 2022: 119).

Different types of trolls are distinguished in the literature. The first type is an advanced troll. These people write a certain number of daily comments on various forums and social networking sites. Such persons are characterised by a lack of profile. They have no photos, posts or information, and their accounts are created on the fly. The posts of an advanced troll are characterised by sharp wording and attack. The second, more complex type is the mole troll. Their profiles are filled out, and the comments are thoughtful and urge discussion. Such persons present themselves as thinking in an avant-garde manner, not being manipulated by generally imposed rules. Another type is the anti-troll, the most complex form of activity, which is difficult to decipher and easily draws people into discussion. Such persons mitigate the dispute of two different sides by polemicising with both sides. There is also a lamer troll. This user knows little about the subject but tries at all costs to prove otherwise. Such persons often discuss plenty and pretend to be professional, using obvious very obvious phrases. Consequently, when dealing with experienced users, they get exposed and are most often blocked by the site moderator (Nowikowska, 2021: 193).

In science, it is assumed that a troll is characterised by the following behaviours: 1) defending a view or an idea, regardless of the statement's veracity; 2) asking questions unrelated to a specific discussion group to make other discussants nervous; 3) repeatedly asking the same questions that have been answered to create forum confusion; 4) not admitting to being wrong; 5) contradicting one's own theses, incompetently leading the discussion, lacking coherent speech; and 6) using personal attacks (Nowikowska, 2021: 194).

These types of actions cannot be effective with sporadic postings. Hence, individual organisations, companies, and governments employ deliberate actions using Internet trolls. These people (i.e., trolls) are paid to place comments and posts on forums. They work 12 hours a day, create around150–200 comments and have several accounts. With 400 people making 200 comments per day, this adds up to about 80,000 daily posts. Such activities, which involve a large group of hired persons influencing public opinion by

multiplying information according to the client's guidelines, can be successful and demonstrate the effectiveness of disinformation. Knowing the characteristics of trolls, an Internet user can detect and defend against them by ignoring their comments and not responding to their taunts. The most important task falls on the administrator or moderator of the website, who should skilfully filter statements and block trolled content (Chałubińska-Jentkiewicz, Nowikowska, 2022: 120).

There is a specific phrase that has been coined in the online community, i.e., "Do not feed the troll", warning against interacting with a trolling person. Hence, the information conveyed by so-called trolls represents a form of disinformation. Contrasting the phenomenon of trolling with "reliable information", it should be noted that information is *"*a set of figures describing objects of any nature, contained in a specific message and expressed in such a form that it allows a specific object, to which it has reached, to take a stance on the situation created by it and to take appropriate actions" (Chałubińska-Jentkiewicz, Nowikowska, 2022: 120). In this aspect, it is possible to demonstrate the primary role of information in shaping national security. In 2002, Osama bin Laden wrote in a letter to the Taliban leader that "Obviously communication in war in this century is one of the most powerful methods of combat. In fact, its ratio may reach 90% of the total preparation for battle" (Paczuszka, 2018: 92). Nowadays, information is seen as an effective tool of warfare, supporting or even replacing existing forms of military confrontation, and a decisive factor for achieving success in future armed conflicts (Batorowska, 2017: 9). The above means that a developmental feature of modern civilisations that has come to the fore is the increasing role of information, as well as disinformation. This is the result of the information revolution, which has brought the world into the era of the information society, where information is the primary product and knowledge the essential wealth (Fehler, 2016: 25). Consequently, the importance of information security needs to be systematically raised.

5 Deepfake

Just a few years ago, the pinnacle of disinformation was disseminating false images and texts. However, the development of modern technology has led to new forms of disseminating false information being created. Nowadays, with the aid of artificial intelligence, it is possible to create a fake video, the so-called deepfake. The video involves replacing the face or body of a specific person with any other character. Consequently, it is possible to change their speech and body movements. The term *deepfake* first appeared in 2017. It was the pseudonym of a user who, with the aid of artificial intelligence, created and published pornographic videos using images of celebrities (Chałubińska-Jentkiewicz 2023: 279). New technologies make us all smarter – should we then worry about combining existing values and the pervasive dominance of technology? It becomes important to ask questions related to the security of development and exploitation of this area. These questions are mainly related to the transformation of citizens into e-citizens, the divergence of interests between market and political

stakeholders and the political scene. The most important issue to be resolved is who decides what is good and what is bad, i.e., what is legal and why (Kerikma, Rull 2016: 13–14).

The concept of deepfake was only coined in 2017, originating from the pseudonym of a Reddit user who published pornographic videos with the faces of porn actresses/actresses swapped for those of celebrities (https://www.vice.com/en_us/article/bjye8a/reddit-fake-porn-app-daisy-ridley [accessed on: 10/01/2020]). It is hard to come up with a uniform definition of the word. .The phenomenon is so new that there is no uniform definition. The most accurate yet careful discussion of the meaning of "deepfake" is provided by Merriam-Webster, also known as the Encyclopaedia Britannica. The authors of the article "Words We're Watching: Deepfake" point to several definitions of the word *deepfake* that have appeared in such media as: *The Independent*, *The Guardian* and *The Washington Post*, and present their own definition based on these examples. The term *deepfake* is usually used to describe a video that has been edited using an algorithm (machine learning) to substitute a person in the original video for someone else (in particular, a public figure) in such a way as to make the video appear authentic.

Based on the above definitions, it may be questionable what exactly the difference between deepfake and fake news is. The purpose of both is to mislead the viewer about the facts. Fake news is a false article in which the author claims that Barack Obama is a Muslim ("Barack Obama is a secret Muslim" https://www.newsweek.com/guide-conspiracy-theories-75003 [accessed on: 10.01.2020]) while deepfake describes a video that someone has modified in such a way as to have Barack Obama praying on it turned towards Mecca.

The key difference between deepfake and fake news is the *de facto* level of "fakeness". Fake news is, after all, the work of a third party, which at best describes some untrue event or distorted statement by a person in a public position. Whereas deepfake directly depicts a non-existent situation/statement. In many texts, you can read that deepfake is a stronger version of fake news. Deepfake does not only involve the substitution of an image but may also include an audio substituted to imitate the voice of a specific person (https://miroslawmamczur.pl/deepfake-co-to-takiego-i-jak-go-zrobic/ [accessed on:02/06/2022]).

In 2018, experts created an exemplary political video in which Barack Obama called President Donald Trump "stupid". Actually, these words were spoken by director Jordan Peele, and the Obama character was generated from other existing footage. The experiment was aimed to show how artificial intelligence can mess up politics. Deepfake technology carries many risks, as it can be used to manipulate public opinion. As this technology develops, fake videos are becoming increasingly difficult to detect. As defined by techtarget.com33, the term *deepfake* refers to an AI-based way of creating or

altering audiovisual content so that it shows a reality that did not or does not exist (https://whatis.techtarget.com/definition/deepfake [accessed on: 08/01/2022]).

The term also refers to audiovisual material created in this way and comes from a combination of the term *deep learning*, which denotes a subcategory of machine learning used by artificial intelligence to improve voice recognition and natural language processing techniques, and the word *fake*, which as an adjective means "false, artificial, forged, counterfeit", and as a noun has the following meanings in Polish: "podróbka" (imitation), "trik" (trick), "hochsztapler" (fraud), "falsyfikat" (afalsification) or "fałszywka" (forgery) (Chałubińska-Jentkiewicz, 2023: 282).

Deepfake audiovisual material is created by two counteracting artificial intelligence algorithms. The first creates the deepfake videos, and the second decides whether the video is real or fake. Each time the video is deemed fake by the second algorithm, the first algorithm learns how to improve the next video to prevent it from being classified as a deepfake. In this way, the algorithms continually improve the quality of the videos they create, which means that they become increasingly difficult to recognise with the naked eye by viewers of audiovisual content, who are largely unaware that such processes are taking place. In fact, until recently, altering video content in an unnoticeable way was difficult and required specialised skills, making it mainly the domain of secret services. Nowadays, anyone can download deepfake software and create a realistic video. According to Andrea Hickerson, director of the School of Journalism and Mass Communications at the University of South Carolina: "Deepfakes are lies disguised to look like truth. If we take them as truth or evidence, we can easily make false conclusions with potentially disastrous consequences. What happens if a deepfake video portrays a political leader inciting violence or panic? Might other countries be forced to act if the threat was immediate?" (https://www.popularmechanics.com/technology/security/a28691128/deepfake-technology [accessed on: 08/01/2020]).

Marco Rubio, a candidate in the 2016 US presidential election, said in turn, "In the old days, if you wanted to threaten the United States, you needed aircraft carriers, nuclear weapons, and long-range missiles. Today, you just need access to our Internet system, our banking system, and our electrical grid and infrastructure, and increasingly, all you need is the ability to produce a very realistic fake video that could undermine our elections that could throw our country into tremendous crisis internally and weaken us deeply" (https://www.csoonline.com/article/3293002/deepfake-videos-how-and-why- -they-work.html [accessed on: 08/01/2020]).

The problem of deepfakes is growing. Forbes magazine reported that the number of such videos on the Internet reached almost 15,000 in 2019, i.e., an increase of 84% compared to 2018 (https://www.forbes.com/sites/johnbbrandon/2019/10/08/there-are-now15000-deepfake-videos-on-social-media-yes-you-should-worry/#2793aa6a3750 [accessed on:

08/01/2022]). Moreover, progress has a significant impact on other areas of human activity. This includes the perception of the surrounding reality, the justice system and the application of the law. Indeed, according to the study titled "Deepfakes and Cheap Fakes", the relationship between media and truth has never been stable. Approaches to how truth is evidenced and perceived have been changed by its existence in cultural, social, and political structures. The authors of the study also point out that the treatment of visual media as an objective documentation of truth is a 19th-century legal construct (Paris, Donovan, 2019: 34).

Deepfake appears to be a younger, more perfect, more effective, more elaborate and more complex version of fake news. Furthermore, while fake news can emerge from basically any human activity, e.g., as an article, a graphic, a video, a song, a rumour, a book, a brochure, an organised event, a meeting, a happening, etc., deepfake most often takes the form of an audio-visual recording created by competing artificial intelligence algorithms that apply machine learning principles and techniques. Deep fake, as a much more technologically advanced product, with the development of artificial intelligence and the information sector in general, has many more perspectives for development and may evolve into more sophisticated and technologically advanced forms of influencing people's behaviour and attitudes (Chałubińska-Jentkiewicz, 2023: 284).

6 Image manipulation

The greatest emotions are triggered by the visual experience. The manipulation of photography is nowadays a widespread phenomenon. Fake news in the form of photographs or graphics has a much stronger effect on the viewer than plain text. The "power of credibility" contained in the image makes the lie more effective and thus more dangerous (Santori, 2009: 49). A commentator in England presented two lists of his favourite films separately on radio, press and television – one was true, and the other evidently untrue. Forty-thousand people were asked which list they thought was true. Radio listeners fared much better in this survey (over 73% answered correctly) than TV viewers (52% correct answers). The conclusion drawn was that viewers are far less critical of false content presented to them and have much less developed symbolic thinking (Santori, 2009: 49).

The more audio-visual stimuli, the more difficult it is to distinguish truth from falsehood – not a very edifying conclusion since we live in an age of dominance of audiovisual messages. Photographs are often used deliberately in the wrong context. The whole setting – the description of the photograph or the article – is incompatible with what the image actually shows. An example is a 2009 picture of a wedding ceremony of Muslims. It shows several-year-old girls who were cousins of the brides getting married at the time and were recognised as brides by everyone. The website Fronda.pl decided to use the photograph as an illustration for an interview with Rev. Prof Paweł Bortkiewicz to illustrate the problem of very young people getting married in Islam. The photo circulated

for a long time in Facebook discussion groups and also repeatedly displayed by right-wing media to prove that Islam has a bigger problem with paedophilia than the Catholic Church (https://konkret24.tvn24.pl/swiat,109/znany-fejk-wyk-used-w-dyskusji-o-filmie-braci-sekielskich,936465.html [accessed on: 08/01/2020]).

Another type of graphic fake news can also be a graphic made using pre-existing photographs. Graphics nowadays is a very rapidly developing area. It is sometimes difficult to distinguish a real photo from an image glued together by the hand of a graphic designer. To sum up, with a wealth of ready-made photographs, it is possible to "conjure up" the image of reality we need. The 2017 Gazeta Polska cover story purported to confirm Jarosław Kaczyński's words about refugees being carriers of dangerous diseases which would lead to an epidemic in Poland. The cover was linked to an article about an alleged epidemiological crisis in EU countries caused by providing aid to immigrants from Africa and the Middle East. This photomontage used two 2016 photos by Rafał Wojczal from his stay at the Al Khazer refugee camp. The third photo was taken by another photojournalist in 2007 in Afghanistan during humanitarian aid organised by Polish soldiers. None of the subjects in this photo were refugees (https://wyborcza.pl/7,75398,22154511,-jak-gazeta-polska-uzyla-ludzi-fotomontaz-o-uchodzcach-w.html [accessed on: 08/01/2020]).

Another fake news photograph involved a complete colour change. The image was "constructed" in 2016 from a 2013 photo showing Pope Francis holding the flag of Argentina in front of St Sebastián's Cathedral in Rio de Janeiro. More colours were added to the flag to make it look like a symbol of the LGBT community (https://konkret24.tvn24.pl/swiat,109/czy-papiez-machal-teczowa-flaga-lgbt,933477.html [accessed on: 08/01/2020]).

To summarise, it can be said that creating fake news through videos is rare. The production of such content is dependent on the financial resources at hand. It is more labour-intensive and often requires expertise. It is also much more difficult than simply posting on Facebook or Twitter. An initiative to combat the scourge of fake news has been introduced by the news service BBC, which has established its internal department to check the veracity of information found on the Internet, and Google has donated £150,000 to fact-checking organisations (Chałubińska-Jentkiewicz, 2023: 287).

7 Fact-checking

According to Eurostat surveys, in 2021, 47% of all people aged 16–74 in the EU saw untrue or doubtful information on news websites or social media during the three months before the survey. However, only around a quarter (23%) verified the truthfulness of the information or content. This information comes from data on ICT usage in households and by individuals published by Eurostat. The idea of fact-checking was born in the United States in 1995 with the creation of Snopes.com, a website dedicated to exposing

fake news. It is the oldest fact-checking site of this kind and is highly valued by journalists. The activities of fact-checking centres can be considered one of the attempts to counter disinformation and information warfare globally (http://demagog.org.pl/krotka-historia-fact-checking/ [accessed on: 10/06/2018]). Their main objective is to tackle false content disseminated through the mass media, mainly online. To this end, using their algorithms, these centres verify and thus control the media content published to present unmasked fake news to the public. The share of people aged 16–74 who verified information found on online news sites or social media in the previous three months was the largest in the Netherlands (45%), followed by Luxembourg (41%) and Ireland (39%). However, the smallest share was recorded in Lithuania (11%), followed by Romania (12%) and Poland (16%) (source: EUROSTAT, ec.europa.eu.). In the EU, people aged 16–74 predominantly checked if the information was truthful by verifying its sources or finding other information online (20%). People also checked information by discussing it with other persons offline or using sources not found on the Internet (12%). The least popular method was checking by following or participating in an online discussion regarding the information (7%).

Undoubtedly, the popularity of fact-checking centres globally increased during the Russian-Ukrainian conflict and after the annexation of Crimea. An example is the already mentioned Ukrainian project, StopFake.org (https://www.stopfake.org/pl/strona-glowna/ [accessed on: 10/06/2022]). It was started by lecturers, graduates, and students of the Mohyla School of Journalism, and attendees in the course for journalists and editors of the Digital Future of Journalism. In 2017, thanks to funding from the Ministry of Foreign Affairs, the Polish version of StopFake.org became fully functional (as the eleventh language version). The StopFake.org project is regularly active on social media such as X (Twitter), Facebook, Vkontakte, Google+, and YouTube. The project's activities focus on publicising a ranking of fake news, e.g., Top 10 absurd fake news, Top 10 fake news about Crimea, etc.

Among the best-known foreign fact-checkers is another Ukrainian non-governmental project called "Information Resistance" [Ukrainian «Інформаційнийспротив», abbreviated – «ІС», Russian. «ИнформационноеСопротивление», abbreviated – «ИС», English. «Information Resistance», abbreviated – «IR»]. It was set up to tackle external threats occurring in the information space, mostly in military, economic, and energy areas, as well as in Ukraine's information security. The project was launched on 2 March 2014, the day of Russia's incursion into Crimea. "Information Resistance" was initiated by the Ukrainian NGO Centre for Military-Political Research (Kyiv) ("Центрвоеннополитическихисследований" г. Киев) (http://cmps.org.ua/ru [accessed on: 10/06/2022]). The Centre for Military-Political Research on its official website states that: "it functions as an independent social organisation that began its activities in September 2008, immediately after the end of the war waged by Russia against Georgia". The authors of the project emphasise that "already then, a group of Ukrainian activists was the first to draw public attention to the influence of Russian propaganda,

manipulation of information in the broad media: radio, press, Internet, leaflets, seeing this as a planned information operation against Ukraine" (http://cmps.org.ua/ru [accessed on: 10/06/2022]). According to the declaration of the founders of this organisation, the credibility of their fact-checking activities lies "in reliable verification of emerging information based on at least two, usually three independent sources. If the information is highly controversial, specialists cite the opinions of witnesses and participants in the events. Any analysis or report is the work of many people". The authors of "Information Resistance" also declare that "information comes not only from their verified sources (personal contacts, people they know) but also from external sources and if new, unverified content is received, analysts apply their own verification algorithm". The creators of the project assure that "only verified information and proven facts are published on the site and that they cooperate with Ukrainian and foreign experts of non-governmental and state structures, as well as experts of international organisations" (http://sprotyv.info [accessed: 10/06/2022]).

In addition to a general section on news from the country, the site contains three area-oriented sections: Kharkiv Information Resistance, Donbas Information Resistance, and South Information Resistance (http://sprotyv.info [accessed on: 10/06/2022]). The Ukrainian fact-checking project is divided into analytical sections. They comprise:

- The Alpha Section (Секция Alpha) which deals with the analysis of information received from sources of Ukrainian state structures;
- The Bravo Section which deals with collecting information on terrorist groups, news on corruption, and state power ministries, and observes the process of purchasing ammunition;
- The Delta Section which analyses external threats to Ukraine, including the Russian Federation;
- The Echo Section which analyses what economic effects military actions have in Ukraine and Russia;
- The Whiskey Section which obtains information from sources in diplomatic circles, including Brussels, the seat of the European Union;
- The Foxtrot Section which analyses information received from so-called emissaries in the combat zone; and
- The Charlie Section which brought together local intelligence (composition only concerned Crimean nationals).

After the Federal Security Service of the Russian Federation began to operate vigorously on the peninsula in mid-March 2014, this Section was closed for the sake of the safety of these people. "Information Resistance" publishes the Top 10 fake news stories of Russian propaganda weekly. For instance, on 22 February 2018, they purged from fiction an article posted on the "Антифашист" ("Anti-Fascist") website about the visit of German MP, representative of the Die Linke party, Andreas Maurer, to the conflict zone of the DRL (Donetsk People's Republic). According to the information provided by the website, the German politician stated that no Russians were stationed on the spot – he did not see

Russian troops (neither Russian soldiers nor Russian military equipment) there. In the opinion of the analysts of "Information Resistance", it is a fact that, firstly, the German party Die Linke is known worldwide for its pro-Russian position, and, secondly, evidence of Russian military involvement in the occupation of Crimea and the conflict in the Donbas is widely available through the websites of analytical groups such as Груз 200, Информнапалм and many others (http://sprotyv.info/ru/search/node/Andreas%20Maurer [accessed on: 14/05/2022]).

The above fake news was disseminated simultaneously through websites, video and TV. Namely, the visit to the conflict zone of the aforementioned Maurer was reported at around the same time by one of the so-called propaganda tubes actively operating in Poland, the Sputnik Polska portal. The portal posted a video of the German MP visiting the A. Norkin's show entitled The Meeting Place. After Maurer's statement, there was a fight during the show's broadcast. More specifically, he stated that the Ukrainian army was responsible for the fighting in eastern Ukraine and the deaths of hundreds of children. This sparked an outcry from Ukrainian political scientist Dmitry Suvorov, who saw Maurer's words as another slander against Ukraine.

In addition, as part of its global cooperation with the media, the International Fact-Checking Network (IFCN) has developed an EU fact-checking website, FactCheckEU.info, bringing together European signatories to the IFCN Code of Principles to counter disinformation in the European Union on a continental scale ahead of the European Parliament elections in May 2019. The European Commission co-finances (with the European Parliament) independent projects in media freedom and pluralism. These projects, among other activities, monitor threats to media pluralism across Europe, create maps of media freedom violations, provide funding for cross-border investigative journalism and support journalists at risk. According to the European Commission, around €40 million (approx. US$44 million) has been invested in such projects since January 2019. In addition, the European Commission has proposed a budget of €61 million (approx. US$68 million) from 2021 to 2027 for the Creative Europe programme, which also supports the audiovisual sectors in Europe in tackling disinformation.

In the 20th century, after the end of the Second World War, communication developed intensively. Since then, the media, or mass media, have significantly influenced the functioning and shaping of society in democratic states. It is commonly believed that the media are the guardians of democracy and the rule of law, thus exercising a monitoring function. It should be added, however, that the influence of the media on the shaping of society and its members can be both positive and negative. Nowadays, new media is the central source of information for mass audiences. We use the Internet to find information and all kinds of communication. Both young and mature recipients use the Internet for entertainment purposes, thus significantly reducing or abandoning the use of print and electronic media. The growing popularity of multimedia should make us reflect. In an era

without social media such as X (Twitter) and Facebook, without Internet sites, the traditional media were responsible for filtering news and describing reality. The responsibility for the information published rested with the journalist or a specific editorial office. The issue of so-called alternative facts, post-truth and fake news did not exist. Today, anyone can post content on the Internet. As a result of unrestricted access to information, haste and anonymity, these messages are often of low quality and lack credibility. Moreover, due to the dominant role of the new media and the occurrence of the aforementioned phenomena, the media have started to construct reality rather than reflect it, as was their original purpose. All the phenomena described above introduce chaos, which may threaten democracy. The threat may concern the sense of security of individual people and, very realistically, the security of the state (Chałubińska-Jentkiewicz, Nowikowska, 2022: 133).

References:

Aral, S., Roy, D. & Vosoughi, S. (2018) *The spread of true and false news online* (Massachusetts: Institute of Technology).

Batorowska, H. (2017) Bezpieczeństwo informacyjne w dyskursie naukowym – kierunki badań, In: Batorowska, H. & Musiał, E. (eds.) *Bezpieczeństwo informacyjne w dyskursie naukowym* (Kraków: Uniwersytet Pedagogiczny w Krakowie, Instytut Bezpieczeństwa i Edukacji Obywatelskiej Katedra Kultury Informacyjnej i Zarządzania Informacją), pp. 9-28.

Chałubińska-Jentkiewicz, K. (2023) *Prawne granice dezinformacji w środkach społecznego przekazu. Między wolnością a bezpieczeństwem* (Toruń: Wydawnictwo Adam Marszałek).

Chałubińska-Jentkiewicz, K. & Karpiuk, M. (2015) *Prawo nowych technologii* (Warszawa: WoltersKluwer Polska).

Chałubińska-Jentkiewicz, K. & Nowikowska, M. (2020) *Ochrona informacji w cyberprzestrzeni* (Warszawa: Akademia Sztuki Wojennej).

Chałubińska-Jentkiewicz, K. & Nowikowska, M. (2022) *Prawo mediów* (Warszawa: C.H. Beck).

Fehler, W. (2016) O pojęciu bezpieczeństwa informacyjnego, In: Kubiak, M. & Topolewski, S. (eds.) *Bezpieczeństwo informacyjne w XXI w.* (Siedlce: Pracownia Wydawnicza Wydziału Humanistycznego Uniwersytetu Przyrodniczo-Humanistycznego w Siedlcach), pp. 25-32.

Formicki, T. (2017) *Walka informacyjna i zarzadzanie przekazem informacyjnym. Antologia wybranych analiz i artykułów z lat 2008-2017* (Warszawa: Fundacja Instytut Informacji).

Jachyra, D. (2011) Trollowanie – antyspołeczne zachowania w Internecie, sposoby wykrywania i obrony, *Zeszyty Naukowe Uniwersytetu Szczecińskiego*, (28), pp. 253-261.

Kerikma, T. & Rull, A. (2016) *The Future of Law and eTechnologies* (Switzerland: Springer International Publishing).

Keyes, R. (2022) *Czas postprawdy Nieszczerość i oszustwa w codziennym życiu* (Warszawa: Wydawnictwo Naukowe PWN).

Krajowa Rada Radiofonii i Telewizji (2020) F*ake news – dezinformacja online. Próby przeciwdziałania tym zjawiskom z perspektywy instytucji międzynarodowych oraz wybranych państw UE, w tym Polski* (Warszawa), available at: htpps:// www.gov.pl/web/krrit/fake-news--dezinformacja-online (February 27, 2024).

Nowikowska, M. (2020) *Granice dozwolonej krytyki działalności prasowej osób pełniących funkcje publiczne* (Warszawa: Wydawnictwo C.H. Beck).

Nowikowska, M. (2021) Zjawisko „trollingu" w Internecie, In: Chałubińska-Jentkiewicz, K., Nowikowska, M. & Wąskowski, K. (eds.) *Media w erze cyfrowej. Wyzwania zagrożenia* (Warszawa: Wydawnictwo Wolters Kluwer), pp. 189-203.

Nowikowska, M. (2023) SYOPS as an element of information warfare, In: Chałubińska-Jentkiewicz, K. & Evsyukowa, O. (eds.) *Information disinformation cybersecurity* (Toruń: Wydawnictwo Adam Marszałek), pp. 163-170.

Paczuszka, A. (2019) *Rosyjskie służby specjalne. Czyli jak rozbroić państwo* (Warszawa: Fundacja Wyszechradzka).

Palczewski, M. (2019) Fake news w polityce. Studia przypadków, *Mediatization Studies*, 3, pp. 137-150.

Paris, B. & Donovan, J. (2019) Deepfakes and Cheap Fakes. The Manipulation of Audio and Visual Evidence, *Data & Society*, available at: https://datasociety.net /library/deepfakes-and-cheap-fakes (February 17, 2024).

Poovey, M. (1998) *A History of the Modern Fact: Problems of Knowledge in the Sciences of Wealth and Society* (Chicago: University of Chicago Press).

Sartori, G. (2009) *Homo videns. Telewizja i postmyślenie* (Warszawa: Wydawnictwa Uniwersytetu Warszawskiego).

Tesich, S. (1992) A Government of Lies, *The Nation*, No. 12, pp. 12-15.

Wasiuta, O. & Wasiuta, S. (2017) Wojna informacyjna zagrożeniem dla bezpieczeństwa ludzkości, In: Batorowska, H. (ed.) *Walka informacyjna Uwarunkowania – Incydenty – Wyzwania. Monografia poświecona Profesorowi Zbigniewowi Kwiasowskiemu z okazji Jubileuszu 50-lecia pracy zawodowej* (Kraków: Uniwersytet Pedagogiczny w Krakowie Instytut Bezpieczeństwa i Edukacji Obywatelskiej Katedra Kultury Informacyjnej i Zarządzania Informacją), pp. 71-90.

Weingarten, M. (2005) *The Gang That Wouldn't Write Straight - Wolfe, Thompson, Didion, and the New Journalism Revolution* (New York: Crown Publishers).

Chapter V

Content Blocking in Light of the Polish Broadcasting Act and the Digital Services Act (DSA) – Comments on the Mutual Relationship of the Acts

GRZEGORZ TYLEC

Abstract The article analyzes the legal regulations: Regulation (EU) 2022/2065 of the European Parliament and Council (Digital Services Act) of October 19, 2022, and the Polish Broadcasting Act in its version amended on August 11, 2021, which introduced changes implementing the provisions of Directive 2010/13/EU on audiovisual media services. This comparison was made because reading of these legal acts may lead to the conclusion that the provisions of these different legal instruments overlap and regulate the same matter, namely the activities of online platforms and video platforms providing intermediary internet services. Therefore, it is necessary to distinguish between these legal regulations and establish their mutual relationship. The main conclusion from the analysis is that, despite the fact that the Directive on audiovisual media services, along with the Polish Broadcasting Act constitutes lex specialis in relation to the Digital Services Act, in practice, the latter will largely shape the functioning of modern internet media and will do so on the same terms for all EU countries.

Keywords: • Digital Services Act • Polish Broadcasting Act • Directive 2010/13/EU on audiovisual media services • intermediary internet services • online platforms

CORRESPONDENCE ADDRESS: Grzegorz Tylec, Ph.D., Associate Professor, Head of the Department of Language, Rhetoric and Media Law, John Paul II Catholic University of Lublin, Faculty of Social Sciences, Institute of Journalism and Management, Al. Racławickie 14, 20-950 Lublin, Poland, e-mail: grzegorz.tylec@kul.pl, ORCID: 0000-0003-2016-4523.

https://doi.org/10.4335/2024.2.5 ISBN 978-961-7124-25-5 (PDF)
Available online at http://www.lex-localis.press.

1 General comments

The issue of the provision of audiovisual digital services within the EU is regulated by Directive 2010/13/EU of the European Parliament and of the Council of 10 March 2010 on the coordination of certain provisions laid down by law, regulation or administrative action in Member States concerning the provision of audiovisual media services (Audiovisual Media Services Directive) (OJ EU L 95, 15.04.2010, p. 1 et seq., hereinafter: "Directive 2010/13/EU"). However, due to the significant technological changes that have taken place in the media services market, the original version of this Directive was modified by Directive (EU) 2018/1808 of the European Parliament and of the Council amending Directive 2010/13/EU on the coordination of certain provisions laid down by law, regulation or administrative action in Member States concerning the provision of audiovisual media services (OJ EU L 303, 15.04.2010, p. 69 et seq., hereinafter: "Directive 2018/1808"). The content of the aforementioned legal acts was implemented into the Polish legal order into the content of the Broadcasting Act of 29 December 1992 (Journal of Laws of 2022, item 1722, hereinafter: "the BA").

The change in the existing legal arrangements made by Directive 2018/1808 was prompted, as indicated in its first recital, by the increased importance of new types of content, such as video clips and various types of user-generated programmes. It was noted that video-sharing platforms and social media services deliver a substantial part of audiovisual content. This can be referred, for instance, to the channels offered on the YouTube platform. The same applies to platforms and services permitting the sharing of audiovisual content (such as Facebook/Meta or TikTok). According to the fourth recital of Directive 2018/1808, these new forms of communication, which have already developed after the adoption of Directive 2010/13/EU, should be covered by Directive 2010/13/EU as long as they can compete for the same audiences and revenues as audiovisual media services. Furthermore, they also have a considerable impact in that they "facilitate the possibility for users to shape and influence the opinions of other users", and they have as their main, and not merely incidental, purpose the provision of audiovisual content of an informative, educational, entertaining nature (Recitals 4 and 5 of the preamble to Directive 2018/1808) (van Drunen, 2020:165). In general, the principal purpose of Directive 2010/13/EU is not to regulate the operation of social media services as these, in principle, serve as a tool for communication between users. In certain situations, they can perform similar functions to traditional media services if adapted appropriately. Within such services, problems may arise with the presence of violence, hate speech and content that is inappropriate for children. Hence, their inclusion in the services regulated by Directive 2010/13/EU should be assessed as justified (Kuklis, 2020: 95).

At the same time, Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC, i.e., the so-called Digital Services Act, was passed on 19 October

2022 (OJ EU.L.2022.277.1, 2022.10.27, hereinafter: "the DSA"), which will take effect on 17 February 2024, except that providers of very large online platforms and very large search engines will have to comply with their obligations under the Act before then. The DSA aims directly to create a safe, predictable and trusted online environment that facilitates innovation and where the fundamental rights enshrined in the Charter of Fundamental Rights of the European Union are effectively protected. The DSA is thus the EU's next major step in regulating the internal market for digital services, following the adoption of the E-Commerce Directive, which has so far been the vital legal regulation here (Buri, Hoboken, 2021: 361).

When analysing the legal regulations of the Polish Broadcasting Act in the version after its amendment of 11 August 2021, caused by the implementation of Directive 2018/1808, and the legal regulations of the DSA, one may have the impression that the legal regulations of these two legal acts refer to the same sphere of the digital services market, which is the activity of online video platforms providing intermediation services. Given the above, there is a need to delineate the scope of these two legal acts and determine their mutual relationship to each other. Thus, the research purpose of this article is to delineate the material scope and to determine the mutual relationship between the BA and the DSA, insofar as they relate to the activities of online video platforms providing intermediation services. This is because, in practice, it is unclear to what extent the activities of video platforms providing intermediation will be governed by the BA and to what extent by the DSA. Will the DSA apply in practice to the activities of traditional media providing their media services on the Internet?

2 The Broadcasting Act as *lex specialis* in relation to the provisions of the Digital Services Act

Referring to the research question outlined above, it should be pointed out that the EU legislator, in creating the DSA regulations, assumed that the legal regulations contained in this legal act would apply only if the Audiovisual Media Services Directive does not regulate an issue. Article 2(4) of the DSA provides that "This Regulation is without prejudice to the rules laid down by other Union legal acts regulating other aspects of the provision of intermediary services in the internal market or specifying and complementing this Regulation, in particular, the following: Directive 2010/13/EU, i.e., the Audiovisual Media Services Directive". When interpreting the aforementioned provision of the BA, it should be stated that the Audiovisual Media Services Directive, and thus the BA, which implements its provisions into the Polish legal order, is to constitute lex specialis to the BA. Such conclusions are also confirmed by one of the recitals of the DSA, where it is stated as follows: "The Regulation is complementary to existing sectoral legislation and does not affect the application of the applicable Union law governing specific aspects of the provision of information society services, which apply as lex specialis. For example, the obligations regarding audiovisual content and audiovisual commercial communications set out in Directive 2010/13/EC, as amended by

Directive (EU) 2018/1808, concerning providers of video-sharing platforms ("the Audiovisual Media Services Directive") will continue to apply. However, this regulation applies to such providers only to the extent that more specific rules set out in the Audiovisual Media Services Directive or other EU legislation do not apply to them".

This clear outline of the relationship between the two legal acts under consideration implies that, in addition to the Broadcasting Act, providers of audiovisual media services on the Internet will also be obliged to comply with the regulations of the DSA, which is in force throughout the EU, in matters not regulated by it. The justification for this conclusion can be found in the subsequent recitals of the regulation, which stipulate that: "This Regulation should complement, yet not affect the application of rules resulting from other acts of Union law regulating certain aspects of the provision of intermediary services, in particular the Audiovisual Media Services Directive". It is, therefore, noteworthy that the provisions of the DSA will apply to matters which are not covered at all, or are only partly covered, by those other legislative acts, as well as to matters where those other acts leave it to the Member States to adopt certain measures at the national level.

As previously indicated in the introduction, both legal acts in question (the DSA and the BA) partly cover the same sphere, i.e., the operation of online platforms that provide intermediary online services. In the BA, this group of entities is referred to as "video-sharing platforms" and in the DSA simply as "online platforms". The material scope of the DSA is defined by Article 1(1), which provides that the Regulation sets out harmonised rules for the provision of intermediary services on the internal market, and it applies to intermediary services provided to service recipients who are established or resident in the Union, irrespective of the place of establishment of the providers of those services. Article 3(G) of the said Act contains a definition of "intermediary services", whereby such services are defined as one of the following information society services:

i) a 'mere conduit' service, consisting of the transmission in a communication network of information provided by a recipient of the service, or the provision of access to a communication network;
ii) a 'caching' service, consisting of the transmission in a communication network of information provided by a recipient of the service, involving the automatic, intermediate and temporary storage of that information, performed for the sole purpose of making more efficient the information's onward transmission to other recipients upon their request;
iii) a 'hosting' service, consisting of the storage of information provided by, and at the request of, a recipient of the service.

The above-indicated material scope of the regulation in question can be compared with the material scope of the BA. According to the current wording of Article 1a, the tasks of radio and television broadcasting, as referred to in the Act, shall be carried out by:

a) providing media services,

b) distributing television programmes, and
c) providing video-sharing platforms.

The tasks of the Polish Regulatory Body for Electronic Media, the National Broadcasting Council, were correlated with this article, where it is indicated that this body safeguards the freedom of speech in radio and television broadcasting, protects the independence of media service providers **and video-sharing platforms providers**, as well as the interests of viewers, listeners and users, and ensures an open and pluralistic radio and television.

From the perspective of the relationship of the legal acts analysed in the body of this article, it is important to note that, in line with the definition contained in Article 4(22a) of the BA, a video-sharing platform is understood as a service that is provided electronically as part of the business activity conducted for this purpose. Hence, the definition formulated in this way implies that if a specific entity operates an online video-sharing platform but does so outside the scope of its business activity, the BA will not apply to this type of activity (Duda-Staworko, 2022: 36). In this case, to the extent not covered by the BA, only the provisions of the DSA will apply.

In addition, it should be noted that the BA has expressly included its application to social media services. This is reflected in the wording of Article 2(2)(6a) of the BA, which provides that the Act does not apply to electronically supplied services allowing content to be shared by their users (social media services), provided that their principal function is not the provision of audiovisual programmes or user-generated videos. This scope of activity of online platforms relating to social media services is, therefore, not regulated by the BA. However, it is covered, in its entirety, by the DSA, even if it is performed by traditional electronic media (e.g., social media of public television).

3 Blocking unlawful content under the Broadcasting Act

The entry into force of the Act of 11 August 2021, amending the Broadcasting Act and the Cinematography Act, resulted in introducing legislation implementing Directive 2018/1808 into the national legal order. This amendment introduced a new chapter 6b entitled "Video-sharing platforms" into the content of the BA. The new provisions in Article 47m contain several information obligations incumbent on video-sharing platforms providers while Article 47n provides for an obligation to apply for registration in the list of video-sharing platforms maintained by the Chairman of the National Broadcasting Council. Subsequent provisions are devoted to prohibitions of posting certain content on video-sharing platforms.

Under Article 47o(1)(1) of the BA, it is prohibited to post on video-sharing platforms any programmes, user-generated videos or other communications that are prejudicial to healthy physical, mental or moral development of minors, in particular, those containing pornographic content or exhibiting gratuitous violence without applying effectual

technical safeguards, as referred to in Article 47p(1). This provision requires video-sharing platform providers to develop and operate effective technical safeguards, including parental control systems or other appropriate measures, to protect minors from access to programmes, user-created videos or other communications that are prejudicial to physical, mental or moral development of minors. The provision also stipulates that video-sharing platforms shall put in place arrangements to enable users to classify their uploaded programmes, user-generated videos or other communications and to apply technical safeguards. The obligation on video-sharing platform providers arising from Article 47p(1), i.e., to apply effective technical safeguards, is aimed at eliminating prohibited content as part of *ex-post* control. The obligations imposed on video-sharing platform providers to use technical safeguards to eliminate unlawful content (so-called content filtering) may not take the form of *ex-ante* control over the content posted by users. This principle arises from Article 28b(3) of Directive 2018/1808.

Article 47o(1), in items (2) and (3), introduces an absolute prohibition on the sharing of video programmes, user-generated videos or other communications:

- that are prejudicial to the healthy physical, mental or moral development of minors;
- that contain incitement to violence or hatred towards a group of people;
- that contain content that may facilitate the commission of a terrorist offence;
- pornographic content with the participation of a minor;
- content inciting to insults to a group of people or an individual;
- content containing prohibited commercial communications, including but not limited to communications containing so-called hidden commercial communications.

Paragraph 2 of the said provision imposed an obligation on platform providers, as entities responsible for how content uploaded to the platform is collated, to apply countermeasures against the publication of unlawful content.

In the context of the obligations of video-sharing platform providers relating to the identification of unlawful or harmful content referred to in Article 47 of the BA, it is worth pointing out the content of Article 47s (1), which states the following: "The provider of a video-sharing platform shall provide transparent and user-friendly mechanisms for the users of that platform to report content published on the video-sharing platform which violates the prohibition laid down in Article 47o". The video-sharing platform provider was obliged to respond to user enquiries immediately, in any case not later than 48 hours after reporting.

An issue worth analysing in the context of the mutual relation of the discussed legal acts is their applicability to blocking the unlawful activity of platform users. A new solution introduced into the Polish legal order by the Act implementing Directive 2018/1808 are rules allowing video-sharing platform providers to block access to content by other users. Under Article 47t of the BA, after requesting the user to remedy the unlawful state within

a set period, the video-sharing platform provider shall prevent access to the programmes posted on the video-sharing platform by its user. Once this is done, the content in the user's account will not be available to the general public. Initially, the content will not be completely removed from the platform but the general public's access to it will be limited only to the user who posted it on the platform. Only through subsequent infringements by the same platform user, the video-sharing platform provider, after requesting the user to remedy the unlawful state within a set period, will be able to block that user's account on the platform for a specified period. The provision of the Act states that the account may be blocked for a period of up to three months in the case of posting, at least twice, programmes, user-generated videos or other communications, despite requesting the user to stop infringing the law, when the content of these materials concerned:

- content that is prejudicial to the healthy physical, mental or moral development of minors, if the video-sharing platform user has not classified it in accordance with the applicable law,
- content in breach of Article 47o (1) (2) and (3),
- content containing prohibited commercial communications (which are in breach of Article 16(1), Article 16b(1) to (3), Article 16c(1), Article 17 and Article 17a or the regulations issued on the basis of Article 47q(2) or, in the absence thereof, which are not marked under the terms and conditions referred to in Article 47r).

In the cases expressly indicated in the wording of Article 47t (3), relating to gross violations of a legal order, the video-sharing platform provider may decide to terminate the user's account permanently. Gross violations of the legal order by the user include the situations described in Article 47o(1)(3), namely:

- publication of content that may facilitate the commission of a terrorist offence;
- pornographic content with the participation of a minor;
- content inciting to insults to a group of people or an individual based on their nationality, ethnic, racial or religious affiliation or lack of religious denomination.

The BA guaranteed every platform user (viewer) the right to report perceived violations (cf. Article 47o.) and imposed an obligation on the platform provider to respond to the person reporting the perceived irregularities. (cf. Article 47s(1).

Similar to the providers of traditional media services, video-sharing platform providers were obliged to store copies of programmes, user-generated videos, commercial communications and other communications made available to the public, for a period of not less than 28 days from the date of their removal from the platform or termination of their availability, and to present them to the President of the National Council upon request.

As for platform users, Directive 2018/1808 does not introduce any specific measures regarding their liability for unlawful content posted on the platforms, except for the sanctions of temporary blocking of the content or termination of the account on the

platform, as described below. Users will, therefore, be held legally liable for the content they publish under the general rules (e.g., for infringement of the Act on Copyright and Related Rights or under the provisions of the Criminal Code for committing the offence of insult or defamation).

4 Blocking unlawful content under the DSA

The DSA does not contain provisions defining what is meant by unlawful content. In this regard, the DSA explicitly states that to determine what content is unlawful, it is necessary to apply the regulations of individual EU Member States and EU law. Recital 12 of the DSA indicates that unlawful content should be understood as any information, irrespective of its form, that under the applicable law is either itself illegal, such as illegal hate speech or terrorist content and unlawful discriminatory content, or relates to illegal activities, such as the sharing of images depicting child sexual abuse, the unlawful non-consensual sharing of private images, online stalking, the sale of non-compliant or counterfeit products, the sale of products or the provision of services in infringement of consumer protection law, the non-authorised use of copyright-protected material or activities involving the violation of consumer protection law.

As far as the DSA is concerned, the literature notes that the content of this legal act has a layered structure consisting of four layers, each regulating a different type of service. The lowest, broadest layer applies to all intermediary services. The next layer consists of obligations applicable only to hosting services, followed by a layer of obligations concerning "online platforms", i.e., entities that, in addition to providing hosting services, store material provided by users and distribute it to the public. The highest layer contains obligations for "very large online platforms" and "very large online search engines". In the lowest and broadest layer, which applies to all intermediary services, the DSA contains provisions on the liability of providers of electronic services. In this regard, the DSA repeats the principles of conditional exclusion of liability for service providers, which were previously found only in the e-commerce directive.

As regards the legal liability of video-sharing platforms for unlawful content posted by their users, nowadays, as before the entry into force of the DSA (based on Articles 12 and 13 of the Act of 18 July 2002 on the Provision of Services by Electronic Means (Journal of Laws 2020, item 344, hereinafter: "the Electronic Services Act"), in the event of unlawful content on the platform, the platform provider is, in principle, not liable for it as long as it has no knowledge of the unlawful nature of the content published by the user (C-236/08 Google France, C-682/18 and C-683/18 YouTube). It should be further pointed out that video-sharing platform providers do not bear editorial responsibility for the content posted on the platforms by users. Providers only put together the content on the platform, and it is somewhat of a rule that they have no knowledge of the unlawful nature of the content published by users (Głowacka, 2016: 185; Klafkowska-Waśniowska, 2014: 130; Klafkowska-Waśniowska, 2016: 45). However, if the video-sharing platform

provider has received information from any source about the unlawful nature of the content distributed on its platform, it is obliged to take action to remove this content. Failure to take the steps prescribed by law will result in the provider's liability for that content (Wilman, 2021: 2190; Wilman, 2022).

Although the DSA does not contain provisions defining the meaning of unlawful content, it does contain specific solutions to help EU Member States better deal with illegal online content. These include rules regarding what the decisions of national judicial or administrative authorities should contain or the obligation for intermediary service providers to take action against certain specific illegal content. Service providers were obliged to implement mechanisms alerting persons suspected of infringing the law. They must deal with them timely, diligently, non-arbitrarily and objectively. Service providers were also obliged to block users who allowed frequent provision of illegal content. Article 20(1) of the DSA states the following: "Online platforms shall suspend, for a reasonable period and after having issued a prior warning, the provision of their services to recipients of the service that frequently provide manifestly illegal content". Similarly, Paragraph 2 of this Article states that the accounts of persons who frequently submit manifestly unfounded notices or complaints will also be suspended.

When an online platform becomes aware of any information giving rise to a suspicion that a serious criminal offence, which may pose a threat to the life or safety of persons, has been or is likely to be committed, it shall immediately inform law enforcement or judicial authorities of the Member States concerned of its suspicions and provide all available information in this regard (cf. Article 21(1)).

The envisaged system of monitoring content by platforms is linked to the obligation of an internal complaint-handling system. The user will have the right to lodge a complaint against decisions of the platform, including:

a) decisions to remove information or disable access to it;
b) decisions to suspend or terminate the provision of the service, in full or in part, towards recipients;
c) decisions to suspend or terminate the account of recipients. (cf. Article 17(1)).

The possibilities for complaints and out-of-court dispute resolution are without prejudice to the users' right to bring an action before the national courts. Judicial redress is not explicitly regulated in the DSA. This means that, in principle, it is an issue that should be regulated in national law.

5 Specific obligations of very large online platforms and very large search engines provided for in the DSA

One of the key obligations under the DSA is to require providers of very large online platforms and very large search engines to assess, and then to address, all systemic risks resulting from the design, operation and use of their services. This has to be done annually. It is a sort of a risk management system – a new solution focusing on problems occurring at the system level, not just on problems pertaining to the individual level. This aims to eliminate not only the effects but mainly the root causes. In drafting the DSA, special attention was also given to dealing with various crises. This Act grants the Commission significant powers regarding providers of very large online platforms and very large search engines, and these providers may be required to do three things:

- to assess whether – and, if so, to what extent and in what way – the operation and use of their services contributes significantly to a severe threat to public safety or public health in the EU,
- to identify and apply measures to prevent, eliminate or reduce such impact; and
- to submit an evaluation report to the Commission on the measures taken.

6 Summary

The comparison shows that although the Audiovisual Media Services Directive and, with it, the BA constitute *lex specialis* to the DSA, this legal act will largely shape how modern online media functions and will do so on the same basis for all EU countries. It can be seen, from the comparison, that the DSA, unlike the BA, will apply to the operation of social media and, in addition, it will also cover the activities of platforms, regardless of whether their providers have the status of business entities. It should be assumed that, even though, formally, the DSA constitutes *lex generalis* to the BA, its provisions will be applied alongside or in parallel with the procedures envisaged in the BA. This is because it is difficult to argue that the applied procedures provided for in the BA would preclude the actions provided for in the DSA.

References:

Buiten, M. (2021) The Digital Services Act: from intermediary liability to platform regulation, *Journal of Intellectual Property, Information Technology and E-Commerce Law*, 12(5), pp. 361-366, http://dx.doi.org/10.2139/ssrn.3876328.

Buri, I. & van Hoboken, J. (2021) The Digital Services Act (DSA) proposal: a critical overview, *DSA Observatory* (Amstardam: Instituut voo Informatierecht), pp. 3-42.

Duda-Staworko, E. (2022) Pojęcie usługi platformy udostępniania wideo w prawie polskim i Unii Europejskiej, *Przegląd Ustawodawstwa Gospodarczego*, (12), pp. 36-46, https://doi.org/10.33226/0137-5490.2022.12.5.

Głowacka, D. (2016) Odpowiedzialność administratorów stron internetowych za treści publikowane przez użytkowników w świetle sprawy Delfi AS v. Estonia, In: Bodnar, A. & Ploszka, A. (eds.) *Wpływ Europejskiej Konwencji Praw Człowieka na funkcjonowanie biznesu* (Warszawa: Wolters Kluwer), pp. 185-212.

Klafkowska-Waśniowska, K. (2014) Nowe formy usług medialnych a przesłanka odpowiedzialności redakcyjnej w dyrektywie o audiowizualnych usługach medialnych, *Zeszyty Naukowe Uniwersytetu Jagiellońskiego. Prace z Prawa Własności Intelektualnej*, 124(2), pp. 112-133.

Klafkowska-Waśniowska, K. (2016) Elektroniczne wersje gazet i czasopism a audiowizualne usługi medialne w prawie Unii Europejskiej i w prawie polskim - glosa do wyroku Trybunału Sprawiedliwości z dnia 21.10.2015 r. w sprawie C-374/14 New Media OnLine przeciwko Bundeskommunikationssenat, *Europejski Przegląd Sądowy*, (8), pp. 45-50.

Kuklis, L. (2020) Media regulation at a distance: video-sharing platforms in AVMS Directive and the future of content regulation, *Media Law*, (2), pp. 95-110.

van Drunen, M.Z. (2020) The post-editorial control era: How EU media law matches platforms organisational control with cooperative responsibility, *The Journal of Media Law*, 12(2), pp. 166-190, https://doi.org/10.1080/17577632.2020.1796067.

van Hoboken, J., Quintas, J., Poort, J. & amp; van Eijck, N. (2019) Hosting intermediary services and illegal content online: an analysis of the scope of Article 14 ECD in light of developments in the online service landscape, *Study for the Commission*, pp. 31-37, available at: https://op.europa.eu/en/publication-detail/-/publication/7779caca-2537-11e9-8d04-01aa75ed71a1/language-en (October 2, 2024).

Wilman, F. (2020) *The responsibility of online intermediaries for illegal user content in the EU and the US* (Cheltenham: Edward Elgar Publishing).

Wilman, F. (2021) The EU' s system of knowledge-based liability for hosting service providers in respect of illegal user content – between the e-Commerce Directive and the Digital Services Act, *Journal of Intellectual Property, Information Technology and E-Commerce Law*, 12(3), pp. 2190-2195.

Wilman, F. (2022) Between preservation and clarification: The evolution of the DSA's liability rules in light of the CJEU's case law, *Verfassungsblog*, available at: https://verfassungsblog.de/dsa-preservation-clarification/ (February 17, 2024).

Woods, L. (2018) Video-sharing platforms in the revised Audiovisual Media Services Directive, *Communications Law*, 23(3), pp. 127-140.

Chapter VI

Disinformation in the Regulations of Selected Countries

KATARZYNA CHAŁUBIŃSKA-JENTKIEWICZ

Abstract Modern democratic legal and political systems, within which public space should serve the free exchange of opinions, are much less able to fulfil their social function as a result of the technological revolution of the 21st century. Media systems have evolved considerably, in which the recipients of messages, who are now also active participants in the social universe of communication, play a fundamental role. The multitude of issues concerning the new sphere of social discourse mobilises legislators at national and regional level to take reasonable care of the legal basis for countering the numerous threats. The main factors disrupting communication are the manipulation and disinformation of messages, deliberately and intentionally formatted for the interests of external actors and by participants introduced at the initiative of external actors. The main research challenge of this article is to analyse the legal arrangements for disinformation in the world. In the light of the current legal solutions, the research objective of the paper should be considered valuable not only from a theoretical, scientific point of view, but also in terms of increasing in practice the possibilities of systemic solutions in the area of threats concerning the security of the individual-citizen in the digital world. The article is based on materials from the author's book entitled 'Legal Limits of Disinformation in Social Media. Between Freedom and Security' (Publisher: Adam Marszałek: Toruń 2023).

Keywords: • cyber attacks • disinformation • freedom of expression • media

CORRESPONDENCE ADDRESS: Katarzyna Chałubińska-Jentkiewicz, Ph.D., Associate Professor, Kozminski University, College of Law, ul. Jagiellonska 57/59, 03-301 Warszawa, Poland, e-mail: kasiachalubinska@gmail.com, ORCID: 0000-0003-0188-5704.

https://doi.org/10.4335/2024.2.6 ISBN 978-961-7124-25-5 (PDF)
Available online at http://www.lex-localis.press.

1 General comments

Disinformation messages are a global problem. Countries are trying to implement their legal and technical solutions to tackle disinformation. As a result – based on different rationales depending on the political system, the nature of governments, and the specificity of the problems related to information disseminated on the Internet – attempts are made to introduce legal regulations regarding responsibility for disinformation activities and mechanisms to influence this type of content and to possibly counteract the dissemination of and access to content deemed to be untrue or to violate certain standards or third-party rights. The selected legal systems presented in the article show the diversity of approaches and the lack of a uniform system of legal solutions, which stems from evident geopolitical, cultural or national differences. However, common and unidirectional regulatory trends can also be observed – especially those that touch on such sensitive elements as fighting against disinformation activities in political advertising and during the election period. Undoubtedly, the events during the elections in the USA and France, and during the referendum in the United Kingdom, indicated the need to move in a regulatory direction – not only in the systems of the countries affected by this type of disinformation, and regardless of the legal culture and administrative and organisational system existing in a given country. It should, therefore, be assumed that the shape of the adopted regulations usually also reflects the specificity of the legal systems and political systems of the jurisdictions in which they were introduced, hence the different regulations of similar problems and the limited transferability of solutions between significantly different jurisdictions (Chałubińska-Jentkiewicz, 2023: 425–425).

2 Australia

The spread of disinformation via the Internet, especially via social media platforms, is recognised as a severe problem in Australia. A global survey conducted in early 2018 showed that trust in the media in the country was at a record low of just 31%, and consumers said they struggled to tell the difference between fake news and facts. Over the past two years, the Australian Government and Parliament have taken several actions relating to protecting democratic systems from interference, including cyber attacks and the spread of disinformation via the Internet. Legislative actions have included strengthening the requirements for authorisation statements for campaign advertisements under election law, with the requirements specifically extended to social media pages and posts. New criminal offences were introduced that concern acts of foreign interference that affect the political or governmental process, the exercise of democratic political rights or duties, or undermine national security. In addition, a new Foreign Influence Transparency Registry has been created, and persons engaging in communications activity in Australia on behalf of a foreign principal, to exert political or governmental influence, must make a statement, also available on social media. Legislation passed in April 2019, following the attacks in Christchurch, New Zealand, requires social media entities to promptly remove abhorrent violent material. Liability for the offence applies to individuals and companies responsible for hosting online content.

Over the past two years, legislative reforms that may affect social media platforms and users have sought to increase the transparency of political advertising, to introduce new offences relating to foreign interference and the sharing of information affecting national security, to establish a registration system and disclosure requirements where communication is made on behalf of a foreign principal, and to impose a new requirement on online companies to remove "abhorrent violent material". The Australian Government introduced the Electoral and Other Legislation Amendment Bill in March 2017. The Bill was enacted in September 2017, and the amendments came into force in March 2018. The Bill aligns election authorisation requirements with modern communication channels, requires all paid election advertising (involving distribution or production) to be authorised, regardless of the source, and ensures that the duty to authorise election and referendum matters rests primarily with those responsible for the decision to provide them, and replaces the current criminal non-compliance regime with a civil penalty regime to be administered by the Australian Electoral Commission. The requirements for the authorisation of political advertising in Australia are contained in XXA Commonwealth Electoral Act 1918 (Cth) (https:/erma. cG753-PB; JSCEM, The 2016 Federal Election: Interim Report on the Authorisation of Voter Communication (Dec. 2016)25).

In April 2019, the Australian Parliament passed legislation establishing new offences in the Criminal Code that require Internet, hosting or content service providers (including social media platforms) to ensure the "prompt removal" of "abhorrent violent material that can be accessed in Australia" and to provide details of such material that was found in Australia to the Australian Federal Police. "Abhorrent violent material" is defined as material that records or transmits abhorrent violent behaviour and is material that "reasonable persons would regard as being, in all circumstances, offensive". It must also be produced by a person who has engaged in violent conduct or who has "aided, abetted, counselled or procured or in any way knowingly participated in abhorrent violent conduct". The offence of failing to remove abhorrent violent material is punishable by imprisonment for up to three years or a fine of up to AU$2.1 million (approximately US$1.47 million) in the case of an individual or a fine of up to AU$10.5 million (approximately US$7.32 million) or 10 per cent of annual turnover, whichever is greater if the offender is a legal entity (Criminal Code Amendment https://perma.cc/UV8K-FHD [accessed on: 21/08/2022]).

3 People's Republic of China

Although the Constitution of the People's Republic of China (PRC or China) declares that citizens enjoy freedom of speech and freedom of the press, these freedoms are not institutionally protected in practice. Freedom House, in its "Freedom in the World 2019" report, states that China is "home to one of the world's most restrictive media environments and its most sophisticated system of censorship, particularly online" (Freedom House, Freedom in the World China Country Report). In November, the

National Radio and Television Administration released new regulations for the country's massive live-streaming industry, which features around 560 million users. The regulations include requirements that platforms notify authorities ahead of celebrity and foreigner appearances, and that they promote accounts embodying core socialist values. The administration also said it would enforce the new regulations during a clean-up campaign in December, during which it would shut down platforms that do not comply (Chiu, 2020). Censors increasingly target "self-media", i.e., the category including independent writers, bloggers, and social media celebrities. Overall, tens of thousands of these accounts have been shut down, delivering a major blow to one of the few remaining avenues for independent and critical news and analysis. The authorities apply pressure on Chinese Internet companies to tightly enforce censorship regulations or risk suspensions, fines, blacklisting, closure, or even criminal prosecution of relevant personnel. Such pressure has intensified under the Cybersecurity Law, which came into force in 2017. (PRC Cybersecurity Law adopted by the Standing Committee of the National People's Congress on 7 November 2016, effective from 1 June 2017. https://perma.cc/3HAP-D6M [accessed on: 21/08/2022]).

From 10 to 17 June 2020, the Cyberspace Administration of China (CAC) suspended the trending topics list for the popular Sina Weibo micro-blogging service, saying messages on the platform had been "disrupting online communication order" and "spreading illegal information". In March 2021, the CAC reportedly ordered Microsoft's LinkedIn to suspend new sign-ups for 30 days and undergo a self-evaluation for not censoring enough content. The company issued a statement on 9 March that it was "working to ensure we remain in compliance with local law".

Despite strict media regulation, disinformation – or what Chinese law often refers to as "gossip" – still seems to permeate the Internet and social media. Internet regulators are said to have received 6.7 million reports of illegal and false information in a single month in July 2018, with many cases coming from Chinese social media platforms Weibo and WeChat. Pursuant to the 1997 State Council Regulation on Computer Information Network and Internet Security, Protection, and Management, it is prohibited to use the Internet to create, repeat, transmit and broadcast information that threatens the implementation of the constitution, laws and administrative regulations inciting to overthrow the government or socialist system, divide the country or threaten national unification, spreading hatred or discrimination against ethnic groups or threatening their unity, spreading rumours or false information, promoting feudalism, obscene material, pornography, gambling, violence, murder, terrorism or supporting criminal activities, violating personal rights, defaming state organisations, as well as any other activity against the constitution, laws and administrative regulations. In contrast, under the 2000 State Council Regulation, websites in China are not permitted to link to foreign news sites or disseminate news from such sites without separate authorisation. In 2016, in the Cybersecurity Law, China criminalised the creation and dissemination of online rumours that threaten economic and social order. In 2017, the Act on the Administration of Internet News Information Service made it mandatory for online news providers to report news

delivered by government-approved news agencies and present it without tampering with or undermining its content. This is to prevent the introduction of messages on social media platforms that do not come from official sources.

In 2018, it was announced that a regulation would be introduced requiring micro-blogging service providers to establish mechanisms to prevent the spread of rumours. On 15 December 2019, the previous scattered regulations were replaced by a new regulation, the Provisions on Governance of the Network Information Content Ecology, issued by the State Internet Information Office, which came into force on 1 March 2020. The addressees of the new regulation are content creators, platforms and Internet users, and it defines prohibited content as illegal, restricted content as harmful and actively promoted content. The actively promoted content should publicise Xi Jinping's thoughts on socialism with Chinese characteristics for the new era, promote the main policies and political thought of the Chinese Communist Party, as well as core socialist values, enhance the international influence of Chinese culture, respond to social needs, teach taste, style and responsibility, proclaim truth, goodness and beauty, and promote unity and stability. Any content that threatens the national unity and national religious policy or gossip that threatens social or economic order, national honour and interests are recognised as illegal content. Online content creators are obliged to take measures to prevent the creation, repetition or publication of negative information, including the use of exaggerated titles, gossip, inappropriate comments about natural disasters, major accidents or other catastrophes, sexual innuendo, sexually related content, fear-inducing content, and things that would push minors into dangerous behaviour or violate social mores. According to the provisions, online platforms are responsible for overseeing all these restrictions. They must set up mechanisms for everything, from reviewing content and comments to real-time checks and handling gossip online. They should appoint a manager for such activities and improve the related staff. The regulation defines content creators as all persons posting any content online. It also places duties on the creators and managers of online groups and forum community sections. Users of information services, online content creators, and online platforms are not allowed to use them for illegal activities. They are also obliged to actively participate in the ecological governance of network information content, regulate illegal and harmful information on the Internet through complaints and reports, and jointly maintain a healthy network ecosystem.

Despite strict regulation of the media and the Internet, disinformation in this country still seems to permeate the Internet and social media in China. China's law prohibits the publication and online transmission of false information disrupting economic or social order. The law also prohibits other information, such as information that may threaten national security, subvert the socialist system or damage the reputation of others. The dissemination of false information that seriously disturbs public order through a news network or other media is punishable by up to seven years in prison. Network operators are obliged to monitor the information disseminated by their users. When a network operator discovers any information that is prohibited by law, it must immediately stop the transmission of the information, delete it, take measures to prevent its spread, keep

appropriate records and report to the relevant government authority. Social media platforms must be licensed to operate in China. Users must provide service providers with their real full names and other identity details. Specific rules have also been established to regulate online news services. For example, when reprinting news, providers of online news services may only reprint what has been published by official state, provincial or other state-designated news organisations.

As of 1 January 2020, new regulations have come into force, prohibiting the publication of deepfake material without proper marking. Any use of them will have to be clearly marked prominently. Otherwise, the dissemination of such information will be treated as a criminal offence (https://www.reuters.com/article/us-china-technology/china-seeks-to-root-out-fake-news-and-deepfakes-with-newonline-content-rules-idUSKBN1Y30VU[accessed on: 11/12/2019]).

Any service that provides information to online users via the Internet is subject to a regulation under which for-profit Internet service providers must obtain a licence to operate from the state authorities. Non-profit providers must also register with government authorities. The regulation requires ISPs to cooperate with government authorities. For example, service providers must keep records of all information published, including their publication dates, as well as information about users, such as their accounts, IP address or domain name, time spent online, etc. Such records must be kept for 60 days and provided to the relevant government authorities upon request. Users are also required to provide service providers with their real full names and details of their identity. Under the Cybersecurity Act, when delivering information publication services or instant messaging services, service providers must require the identity details of users. Service providers are prohibited from providing the relevant services to those users who fail to perform identity authentication. In cases where service providers fail to authenticate users' identities, competent authorities may order them to take corrective action, suspend their operations, close down their websites, revoke their operational permits or business licenses, or impose a fine of RMB 50,000 to RMB 500,000 (approx. US$ 7,500 to US$ 75,000) on service providers and/or a fine of RMB 10,000 to RMB 100,000 (approx. US$ 1,500 to US$ 15,000) on responsible persons.

Tencent, the operator of China's biggest social media platform WeChat, released a January 2019 report regarding its fight against gossip spread online. According to the report, WeChat intercepted over 84,000 pieces of gossip in 2018. In addition, thousands of "articles" were published through WeChat by government authorities in charge of the Internet, public safety, food and drugs.

4 Russian Federation

In the authoritarian political system of the Russian Federation (RF), power is concentrated in the hands of President Vladimir Putin, who brings together around him loyalist security services, a subservient judiciary, a legislature made up of the ruling party and flexible opposition groups, and above all a controlled media environment. An additional aspect of the functioning of the media market is the rampant corruption that thrives on the close links between officials and organised crime groups.

The Government of the Russian Federation recognises information security as an integral part of national security. Two key documents – the Doctrine of Information Security and the 2017–2030 Strategy for the Development of an Information Society in the Russian Federation – set priorities for information security and identify the main threats and ways to counter them. The Constitution of the Russian Federation contains guarantees of freedom of expression, and various aspects of information integrity, including information on election campaigns, are regulated by federal laws such as the Law on Information, the Law on Mass Media and the Law on Basic Guarantees of Electoral Rights. Recently adopted legislation restricts access to information containing fake news or offensive and disrespectful messages regarding the symbols of the Russian Federation, the Constitution and the authorities. The dissemination of prohibited information is punishable by fines and administrative arrest. The Criminal Code of the Russian Federation contains articles providing for various penalties for disseminating defamatory content. Measures to remove prohibited content and restrict access to websites containing proprietary information were introduced in 2019.

The Russian government has created an open register of fake news sites, with the identification of platforms and their authors. The lower house of the Russian legislator plans to study news aggregators to control the distribution of fake news and disinformation. The Internet and social media are widely accessible and reachable for a large part of the Russian population. According to the statistical website Statista, the number of Internet users in Russia has grown steadily over the past six years, reaching one hundred million users in 2019. According to the same source, the majority of the Russian population uses social media. As of 2017, the most popular social networks in the Russian Federation were YouTube (68%) and VKontakte (61%). For the government of the Russian Federation, information security is an inseparable component of overall national security (Statista, 2019, https://perma.cc/NS4X-ZE3X [accessed on: 21/08/2022]).

The Government's Doctrine on Information Security emphasises the importance of regulating the Internet within the borders of the Russian Federation. It considers all content containing extremist ideology, spreading xenophobia, promoting violent changes to the constitutional order or violating the territorial integrity of the Russian Federation to be a security threat. Based on the principles and priorities outlined in the Doctrine, Russia adopted the Strategy for the Development of an Information Society in the Russian

Federation for 2017–2030 (Resolution of the President of the Russian Federation on Approving Information Security Doctrine (5 December 2016) (in Russian), https://perma.cc/4BEK-4M5R [accessed on: 21/08/2022]). One of the declared objectives of the Strategy is to "create a secure information environment based on information resources that contribute to the dissemination of traditional Russian spiritual and moral values". To pursue this objective, it is planned to amend the legal, regulatory and technological systems to protect the information sphere in Russia by blocking access to and removing prohibited resources (Decree of the President of the Russian Federation on the Strategy for the Development of an Information Society in the Russian Federation for 2017–2030, N 203 (9 May 2017), http://pravo.gov.ru (official legal information portal) (in Russian), https://perma.cc/AQ4H-CE79 [accessed on: 21/08/2022]).

In March 2019, Russia adopted two so-called anti-fake news laws that amended the Federal Law on Information. It introduced provisions establishing a procedure for removing information deemed false and providing for punitive measures for the dissemination of fake news. At the same time, the Law on Information and the Code of Administrative Offences were amended with provisions prohibiting the publication on the Internet of content that insults state symbols, the Constitution and the authorities of the Russian Federation. Some provisions of the Criminal Code provide for penalties for disseminating inaccurate, defamatory and false content (Federal Law on Information, Information Technologies and Protection of Information, No. 149-FZ (27 July 2006) https://perma.cc/86PF-DYTH [accessed on: 21/08/2022).

5 France

Two areas are the subject of French regulation: defamation and fake news, on the one hand, and advertising, including political advertising, on the other. Some laws have been in place for a long time but the emergence of social media has created challenges that have prompted the recent adoption of new ones. Freedom of expression is considered a "fundamental freedom" in France. It is protected by the French Constitution, which includes the 1789 Declaration of the Rights of Man and the Citizen. Articles 10 and 11 of the Declaration protect freedom of opinion and expression, describing the "free communication of ideas and opinions" as "one of the most precious rights of man". However, freedom of speech was never intended to be absolute. Unlike the First Amendment to the US Constitution, the 1789 Declaration of the Rights of Man and the Citizen provides for limitations to freedom of expression in the definition itself. On 22 December 2018, President Emmanuel Macron signed a new law against disseminating false information (Law No. 2018–1202 of 22 December 2018 on the fight against the manipulation of information (22 December 2018), https://perma.cc/QH5N-25MC [accessed on: 21/08/2022]). This legislation was adopted in reaction to new methods of disseminating disinformation, the Internet in general and social media in particular. Under this new Law, online platforms are obliged to establish a way for users to flag false information, especially in content promoted by a third party. This method of flagging fake news must be "easily accessible and visible". Furthermore, online platforms are

encouraged to take measures such as improving the transparency of their algorithms, promoting content from press agencies and radio and television services, fighting against accounts that massively disseminate fake information, informing users of the identity of the person(s) or organisation(s) that bought paid content related to "a debate of national Interest", informing users of the nature, origin, and manner of broadcasting content, and educating people about the media and information. Online platforms must provide the Conseil supérieur de l'audiovisuel (CSA) (the National Council on Audiovisual), France's main regulatory agency for radio and television broadcasting, with a yearly statement indicating what measures they took to fight against fake news. The CSA is then expected to publish regular reports on anti-fake news measures taken by online platforms and their effectiveness. Additionally, online platform operators that use algorithms to organise the display of content related to "a debate of national interest" are required to publish statistics on how they work.

For every item of content, online platform operators must specify how often it was accessed directly, through the platform's recommendation, sorting, and referencing algorithms, and through the platform's internal search function. These statistics are to be published online and made accessible to anyone.

Online platform operators must designate a legal representative in France to serve as a point of contact for applying these provisions. Some provisions of this new Law aim to improve transparency for political advertising on the Internet. Specifically, the Law amended the Electoral Code to provide that online platforms with at least five million unique visitors per month must, during the three months preceding the first day of a month during which a national election is scheduled and until the end of that election, provide users with "faithful, clear, and transparent information on the identity" of the person(s) or organisation(s) that bought paid content related to "a debate of national interest". Additionally, during that same timeframe, online platforms are required to give their users "faithful, clear and transparent information on the use of their data in the context of promoted information content related to a debate of national interest", Furthermore, during the same period, online platforms that are paid €100 (approximately US$110) or more per sponsored content must make the payment amount public. Failure to abide by these requirements is punishable by up to one year in jail and a fine of €75,000 (approximately US$83,150).

The new Law also creates a new legal weapon to combat disseminating fake news during an election period. During the three months preceding the first day of an election month and until the end of that election, a judge may order "any proportional and necessary measure" to stop the "deliberate, artificial or automatic and massive" dissemination of fake or misleading information online. A public prosecutor, candidate, political party or coalition, or any person with standing may file the motion, and the court must rule within 48 hours. Additionally, the CSA may suspend the broadcasting license of an operator controlled by or under the influence of a foreign state if, during an election period, if it broadcasts false information that could affect the election results. While this measure is

aimed at radio and television broadcasters, a suspension ordered by the CSA may apply to broadcasts on "any electronic communication service" (i.e., the Internet) and radio and television broadcasting. The CSA may also, after a first warning, withdraw the broadcasting license of a radio or television operator controlled by or under the influence of a foreign state if it broadcasts harmful content. This provision explicitly states that spreading false information to interfere with the proper functioning of institutions should be considered harmful to fundamental national interests. The CSA may, in deciding to withdraw a broadcasting license, consider content that the broadcaster, its subsidiary or parent organisation published on other services, such as the Internet. However, the CSA may not base its decision to withdraw a license entirely on that factor.

A key factor in countering foreign intervention efforts appears to have been the active role of two government agencies: the Commission Nationale de Contrôle de la Campagne Électoraleenvue de l'Élection Présidentielle (CNCCEP) (the National Commission for the Control of the Electoral Campaign for the Presidential Election), and the Agence Nationale de la Sécurité des Systèmesd'Information (ANSSI) (the National Cybersecurity Agency) These agencies worked with the presidential candidates' campaigns to educate them on cybersecurity and warn them of specific threats and attacks.

The law against disseminating false information adopted in December 2018 provides that French public schools should teach students how to navigate online information. These recommendations largely reiterated those set out in Law No. 2018–1202 of 22 December 2018 and include implementing an accessible and visible reporting mechanism, ensuring transparency of algorithms, promoting content from newspapers, news agencies and audiovisual communication services, detecting and countering accounts that massively disseminate false information, ensuring transparency of promoted content and promoting the skill to media and information.

6 Spain

The Spanish legislator aims to introduce the crime of disinformation or the deliberate dissemination of false information through the use of digital global communication platforms, Internet technologies, any computer system or any means of communication or data transmission technology suitable for altering the regular results of election acts, but this applies to the Election Code. The manipulation of political processes through digital media and social media to cause disinformation, either through confusion, by fragmenting and dividing societies, or by breaking down the social fabric and creating an environment conducive to xenophobic politics, is identified as a threat. In addition, the government is working on introducing a rapid alert system (rapid alerts) against fake news so that it can be responded to immediately. For now, Spain will participate in the coordination of the strategy for the denial of fake news. Joining the European strategy is expected to allow rapid action sufficient to detect fake news (European action plan against disinformation).

7 Israel

Cybersecurity is seen by the Israeli government as an important national security interest due to geopolitical considerations. The rapid pace of technological progress in cyberspace has raised particular concerns in recent years about the ability of external and internal actors to manipulate public opinion through spreading disinformation on social media and the impact of this development on democratic governance. Specific concerns about foreign intervention in Israel's general elections were particularly highlighted in the run-up to the elections of 9 April 2019. Except for media reports of Iranian intelligence hacking into the mobile phone of Benny Gantz, Chairman of the Kahol Lavan political alliance, no specific data have been published on incidents of cyber attacks, the spread of false information or other improper online behaviour concerning the Knesset elections.

However, in the end, the biggest threat may come from people trying to manipulate opinions by spreading misleading information online, for example, through fake Facebook profiles. The number of bots – fictitious social media users – can be huge. Bots can be created and maintained for three or four years and activated when the elections start. The challenge is to maintain credibility and public trust in the process. Sometimes, it is enough to block a government website for a few hours to raise public doubts about the purity of the system.

As claimed by Tamir Pardo, Head of the Mossad (Israel's secret intelligence service),"What we've seen so far with respect to bots and the distortion of information is just the tip of the iceberg. It is the greatest threat of recent years, and it threatens the basic values that we share - democracy and the world order created since World War Two" (Ziv, 2019).

Experts say that although protecting critical infrastructure and organisations from cyber attacks is a challenge that should be mastered, the battle for public opinion caused by the spread of disinformation requires more complex treatment. The complexity of finding appropriate legal remedies stems from the need to balance the objective of cybersecurity with constitutional principles such as freedom of expression, the right to privacy, the purity of elections, the principles of transparency and parliamentary oversight of government activities, etc. An additional challenge for securing cyber systems is that legal regulations often lag behind the continuous development of new technologies. Several legislative proposals have been put forward regarding cybersecurity and the specific threats posed by the spread of disinformation. These include a proposal for a law regulating the mission, functions and objectives of the Israel National Cyber Directorate, and its authority to detect and identify cyber attacks on Israel, and to warn and share information about such attacks.

Other proposed laws specifically address transparency requirements for online political advertising and removing foreign-funded and harmful online content. Although the statutory transparency requirements for election propaganda were originally limited to

print advertisements, the Central Elections Committee (CEC) has extended them to online election advertisements ahead of the national elections on 9 April 2019. The CEC also recognised the government's obligation to refrain from publishing misleading information.

Ahead of the 9 April 2019 elections, Facebook blocked anonymous and paid Israeli political ads on its site, whilst Google blocked all advertising options related to segmentation, retargeting and using a list of names by anyone involved in political advertising. Addressing the challenges of disinformation, the CEC for the upcoming 17 September 2019 national elections has posted recommendations for identifying the government's response to disinformation on social media platforms and video clips to clarify its message on the subject.

Cyberthreats to Israeli targets can come from both foreign and domestic sources. The ability to spread disinformation on social media easily and quickly, and thereby to manipulate public trust in national institutions or public opinion on other issues, is considered a growing challenge by Israeli policymakers and experts. However, tackling the spread of disinformation on social media through legal regulation raises serious constitutional, institutional and ethical concerns. Among the technological tools used in the battle for public opinion, experts cited bots, big data, hacking and trolls. Bots can spread countless messages encouraging controversy, hatred and violence in the form of posts or talkbacks to articles published in online newspapers. The use of big data analytics makes it possible to target specific audiences based on political preferences or perceived susceptibility to manipulation, as revealed by a person's record of online activity on Facebook or other networks. Other means of possible online manipulation included the hacking of legitimate accounts, the use of professional paid "talkbackers" (trolls) and the impersonation of innocent forums to recruit followers in order to prepare the infrastructure of followers for the "command day".

Deepfake is a new AI-based technology that facilitates "a combination of 'deep learning' and 'fake news' [and] enables the creation of audio and video of real people saying words they never said or things they never did". Such technology can be used to create fear, the perception of a lack of control and harm to a person's privacy "in ways never thought of before". Most important are the wider social implications of this technology. It is not just the fear of false imitation of political candidates. According to Israeli experts, deepfake technologies lead to an inability to distinguish truth from lies, increasing challenges in explaining reality and the phenomena and processes taking place, and the distrust of ourselves and our ability to determine right and wrong in the world around us. Together, these three threaten the foundations of government, the functioning of institutions and the ability to maintain viable human and social relationships.

As in other technological contexts, there are three ways to deal with the threat of deepfakes. The first is to raise public awareness to identify fakes, first and foremost, by asking questions. The problem is that sometimes the impact on people's awareness

remains even after they realise it is a fake. Moreover, teaching people not to believe anything comes at a great social cost. The second way is to create a cat-and-mouse race between deepfake creators and those who develop identification technologies. A third way is to regulate the development and distribution of deepfake products. The authors suggest that there may be a basis for distinguishing between the regulation of fake news and deepfakes, noting that in the US, social networks are exempt from liability for the content that passes through them and is created to support the growth of the Internet.

Social polarisation, hate speech and fake news have not yet caused lawmakers to revoke the exemption, but deepfake may be a reason to impose such liability. It is worth recalling the words spoken by Mark Zuckerberg, who claimed that Facebook might treat deepfakes differently from fake news. To illustrate the challenge posed by the use of deepfakes, the authors cite the case of Deep Nude, a deepfake app that allows the creation of nude images of women based on their images in clothing, using a machine learning algorithm. After half a million downloads and a server crash, the software was removed by its creator. The Deep Nude story teaches again that there is no need to do good in technology, and the challenge lies in setting moral boundaries. Recently, there have been claims that it is not enough to take ethical considerations into account when creating educational systems, but there are educational systems that do not need to be created at all, even by legal prohibition, against all the difficulties this creates. The creator of the Deep Nude software removed it from the servers, claiming that "the world is not ready yet". For this, we can say that we are thoroughly ready. We just don't want it. Constitutional challenges associated with regulating the dissemination of disinformation concern the impact of regulating the dissemination of information on protecting the freedom of expression and the right to privacy. In addition, regulating cybersecurity at the national level may undermine, for instance, the principles of transparency, parliamentary oversight and equality in elections.

8 Canada

No regulation in Canada expressly prohibits the dissemination of false news, even if it is defamatory. Attempts to address the problem of disinformation must be balanced against the right to freedom of expression protected by Subsection 2(b) of the Canadian Charter of Rights and Freedoms, which states that everyone has the fundamental freedom of "thought, belief, opinion and expression, including freedom of the press and other media of communication". Fundamental rights, including freedom of expression, are subject to Article 1, which allows for "reasonable" limits on these rights. This means that once a Charter right is found to have been infringed, the courts must decide whether the right has been infringed. Section 181 of the Criminal Code of Canada prohibits the dissemination of false news ("Everyone who wilfully publishes a statement, tale or news that he knows is false and that causes or is likely to cause injury or mischief to a public interest is guilty of an indictable offence and liable to imprisonment for a term not exceeding two years").

The Elections Modernisation Act, passed in December 2017 and entirely in force from 13 June 2019, amended the Canada Elections Act (CEA) and other laws to modernise Canada's election law. According to a government news release, "the new legislation is part of a comprehensive plan to safeguard Canadians' trust in our democratic processes and increase participation in democratic activities".

Among the changes included in the Act was a provision that considered it an offence "to make false statements about a candidate to affect election results". In particular, the Act provided that no person or entity shall, with the aim of influencing the results of an election, make or publish, during the election period: a) a false statement that a candidate, a prospective candidate, the leader of a political party or a public figure associated with a political party has committed an offence under an Act of Parliament or a regulation made under such an Act – or under an Act of the legislature of a province or a regulation made under such an Act – or has been charged with or is under investigation for such an offence; or b) a false statement about the citizenship, place of birth, education, professional qualifications or membership in a group or association of a candidate, a prospective candidate, the leader of a political party or a public figure associated with a political party.

The Elections Modernisation Act also aims to prevent foreign interference in the election process regarding paid political advertising through online platforms. Foreigners and foreign entities may not purchase regulated advertising during the election period, currently defined as a maximum of 15 days. Platform operators or owners may be prosecuted (or other compliance or enforcement action may be taken) for knowingly selling election votes. Third parties may not use funds for regulated activities, including election advertising, if the source of the funds is a foreign entity; prohibits foreign third parties from participating in elections and incurring expenses for regulated activities (including partisan advertising expenses) that are undertaken by foreign entities.

The Elections Modernisation Act also imposes requirements on online platforms to improve transparency and integrity of content during elections. Section 319 of the CEA defines an "online platform" as "an Internet site or Internet application whose owner or operator, in the course of their commercial activities, sells, directly or indirectly, advertising space on the site or application to persons or groups". Platforms in this category must maintain a digital register of all regulated advertisements, publishing the register and details of the agents who have authorised the advertisements. Ads must be placed on the register on the day they first appear on the platform, and each ad must be kept on the register for two years after the election. After this period, operators or platform owners must keep the ad information for five years.

9 Norway

A fact-checking initiative called Faktisk was set up in Norway before the 2017 general elections. It was created jointly by two tabloids, *Verdens Gang* and *Dagbladet*, the public broadcaster NRK and the commercial TV channel TV2. Funding for the initiative comes from the owners of the four publishers and the freedom of expression organisation, Fritt Ord. Faktisk checks news appearing in Norwegian media and social media, in public debates and statements by politicians, and follows up on complaints made by the public. The main topics are climate, Norwegian elections and international affairs. Faktisk ranks each submission on a veracity scale of one to five, making it available as text or a short video on its website, through social media platforms such as Facebook and Snapchat, and on television. It uses open formats for these purposes so that other media companies can use its resources. The Faktisk website is one of the most popular in Norway.

Another initiative aimed at civil society in Norway is a fact-checking tool for newspaper readers, called Reader Critic, developed by *Dagbladet*. This system allows readers to report inaccuracies in the newspaper's content and automatically notifies the author. In the first nine months of the Reader Critic programme, *Dagbladet* received 20,000 opinions on 10,000 articles from 5,000 users. The information most often pointed out grammatical errors. However, some more serious errors were also identified.

10 Sweden

In Sweden, there is a focus on cooperation between the public sector and the private (media) sector. A new government-funded cooperative between the public service and the three largest media houses in Sweden (Schibsted, Bonnier and NTM) has been announced. Together, they will develop a digital platform to counter the spread of fake news, an automated news rating service, and an automated tool for checking and personalising facts. Sweden relies on free media. It takes the position that the best protection against fake news is free media that compete with each other and "breathe down each other's neck". The fact that they now collectively decide what fake news is prevents the misinformation passing through the media network from becoming more widespread and legalised.

As Sweden points out, there is a risk the reaction of the media and social networks to fake news will increase distrust as well as become a tool for silencing divergent views.

The line between opinion and information, and between fake news and true news, is extremely difficult to draw. If done wrong, the effort will be transformed from an attempt to prevent the spread of fake news into a tool to prevent the spread of unpopular opinions. In Sweden, there has been an initiative to create an organised control of information, with the media playing a large role. The cooperating media are to individually review information spread on social media from individuals and political authorities. The collected material is then to be presented on a shared website. Carefully reviewing the

data and searching for its source is very time-consuming, so there is a reliance on media cooperation. By combining multiple media in this project, the public can access accurate information. The cooperation between numerous media also means that the correct information can be more easily accessed on social media. In the case of false information regarding the coronavirus, the AFP News Agency publishes daily fact-checking articles regarding it.

Sweden recognises the right to freedom of expression, including online and through using social media platforms. While private entities are free to block inappropriate content, the government neither prohibits using Twitter or fake Twitter accounts nor has it passed legislation allowing the government to block websites or Internet access. It does not regulate opinion-based advertising either. However, Sweden has recognised spreading false information as a criminal offence and obliges the news media to correct such information. Realising that disinformation is a significant global challenge, the Swedish government is in the process of launching a new agency, the Psychological Defence Agency, which will focus on psychological defence and combating disinformation in Sweden. The agency is expected to be launched in 2022. The Swedish Emergency Agency had previously been tasked with making the Swedish population aware of disinformation campaigns and educating them on how to check the veracity of information and was actively involved in this process. Media companies have begun to address disinformation voluntarily. During the 2018 national election cycle, four Swedish public media corporations created a fact-checking website (now discontinued) that allowed members of the public to verify election-related claims. Bots were used in the 2018 elections, but no successful disinformation campaigns were identified. Facebook removed posts that contained false information produced by fake accounts in connection with the 2018 national elections. TV4 initiated rules prohibiting the purchase of political advertising by foreign entities in the weeks leading up to the 2019 EU parliamentary elections. Disinformation continues to be one of Sweden's challenges, from the perspective of defence and civil emergencies. The mass dissemination of disinformation is recognised by the Swedish authorities as a global problem. The risk of future mass dissemination of information in Sweden, especially about elections, is also recognised. Sweden protects the right to freedom of speech as enshrined in its Constitution (Instrument of Government). Further regulation of freedom of expression is contained in two separate constitutional acts, the Law on Freedom of the Press (Tryckfrihetsförordning, TF) and the Basic Law on Freedom of Expression (Yttrandefrihetsgrundlagen, YGL). Sweden introduced the first legislation concerning freedom of the press in 1766.

References:

Chałubińska-Jentkiewicz, K. (2023) *Prawne granice dezinformacji w środkach społecznego przekazu. Między wolnością a bezpieczeństwem* (Toruń: Wydawnictwo Adam Marszałek).

Chiu, K. (2020) China orders live streamers and gift-giving fans to register with real names, *South China Morning Post*, (November 24, 2020), available at: https://www.scmp.com/tech/policy/article/3111177/china-orders-live-streamers-and-gift-giving-fans-register-real-names (August 21, 2022).

Criminal Code Amendment, available at: https://perma.cc/UV8K-FHD (August 21, 2022).

Decree of the President of the Russian Federation on the Strategy for the Development of an Information Society in the Russian Federation for 2017–2030, N 203, (May 9, 2017), available at: http://pravo.gov.ru, https://perma.cc/AQ4H-CE79 (August 21, 2022).

Federal Law on Information, Information Technologies and Protection of Information, No. 149-FZ, (July 27, 2006), available at: https://perma.cc/86PF-DYTH (August 21, 2022).

Law No. 2018–1202 of 22 December 2018 on the fight against the manipulation of information, (December 22, 2018), available at: https://perma.cc/QH5N-25MC (August 21, 2022).

PRC Cybersecurity Law, adopted by the Standing Committee of the National People's Congress on 7 November 2016, effective from 1 June 2017, available at: https://perma.cc/3HAP-D6M (August 21, 2022).

Resolution of the President of the Russian Federation on Approving Information Security Doctrine, (December 5, 2016), available at: https://perma.cc/4BEK-4M5R (August 21, 2022).

Reuters (2019) *China seeks to root out fake news and deepfakes with new online content rules*, available at: https://www.reuters.com/article/us-china-technology/china-seeks-to-root-out-fake-news-and-deepfakes-with-newonline-content-rules-idUSKBN1Y30VU (December 11, 2019).

Statista (2019) *The number of Internet users in Russia*, available at: https://perma.cc/NS4X-ZE3X (August 21, 2022).

The 2016 Federal Election Interim Report on the authorisation of voter communication. Joint Standing Committee on Electoral Matters, available at: https:/erma. cG753-PB; JSCEM (August 21, 2022).

Ziv, A. (2019) Massive Manipulation, Foreign Influence Campaign and Cyber: The Threats to Israel's Election, What's behind the Shin Bet Chief Warning that a 'Foreign Country' Intends to Intervene in the Israeli Election, *Haaretz.com*, (January 9, 2019), available at: https://perma.cc/LE7Y-79SN?type=image (August 21, 2022).

Institute for Local Self-Government Maribor

www.lex-localis.press
info@lex-localis.press