AGENCY FOR ELECTRONIC MEDIA

STUDY

STRENGTHENING RESILIENCE TO DISINFORMATION:

THE STATE OF AFFAIRS AND GUIDELINES FOR ACTION

Dr. Marijana Grbeša Zenzerović
Dr. Iva Nenadić

NATIONAL RECOVERY AND RESILIENCE PLAN 2021 – 2026
INTRODUCING FACT-CHECKING AND ESTABLISHING AN OPEN DATA DISCLOSURE SYSTEM

NATIONAL RECOVERY AND RESILIENCE PLAN 2021 – 2026
INTRODUCING FACT-CHECKING AND ESTABLISHING AN OPEN DATA DISCLOSURE SYSTEM

STUDY

STRENGTHENING RESILIENCE TO DISINFORMATION:

THE STATE OF AFFAIRS AND GUIDELINES FOR ACTION

Dr. Marijana Grbeša Zenzerović
Dr. Iva Nenadić

Authors of the chapter "Computational fact-checking systems":
Matej Gjurković & Dr. Mladen Karan

Editor: **Robert Tomljenović**

Design and layout: **Creatorium**

Publisher: **Agency for Electronic Media**

Second supplemented edition

Zagreb, September 2022

Financira Europska unija
NextGenerationEU

AGENCY FOR ELECTRONIC MEDIA

Republika Hrvatska
Ministarstvo kulture i medija
Republic of Croatia Ministry of Culture and Media

Content

1. Introduction

For the past several years, the problem of spreading online disinformation has been recognised as one of the key challenges for democracy, whereas in the context of the COVID-19 pandemic, it turned out to be a major challenge for health and human lives as well. Disinformation is not a new phenomenon, yet in the modern information environment there is a whole range of new methods, techniques, and technologies of producing and spreading manipulation at an unprecedented speed and with a scope of impact that was not previously possible. On the one hand, the volume and diversity of news and sources has increased, potentially contributing to the democratisation and pluralism of public debate. On the other hand, it seems that the prevailing trend consists in the proliferation of online information, informants, and commentators, which has resulted in an information disruption where it is difficult for citizens to distinguish credible from misleading information. Such an environment may encourage distrust of traditional media and institutions, thus positioning alternative sources of information, whose credibility is rather questionable, as the main "sources of truth" for a significant number of users of social media.

The entire information environment has changed with the emergence of new technologies and actors, new power relations, altered habits of media audiences and new practices of the media themselves. Nowadays, traditional media are becoming increasingly dependent on social media when finding topics and sources, distributing their content and are further conditioned by the challenge of reaching the audience. Online platforms, such as social networks, search engines, news aggregators, messaging apps or video sharing apps, are positioning themselves as central places for gaining various information, including news and information exchange. Thus, platforms are taking on the role that the media have played for a long time. In most cases, they do not produce their own content, but rather transmit content produced by their users, including the media, through algorithmic, not editorial decisions, and driven primarily by commercial rather than public interest. Algorithmic processes of ranking and adjusting visibility of content as higher or lower affect the scope and spread of disinformation, as opposed to the reach of credible sources of information.

In an environment of information disorder, the media are simultaneously part of the problem and victims of such an environment. With online platforms dominating digital advertising, the media are forced to develop alternative business models to ensure sustainability and remain relevant. However, in this process, the media often follow the imperative of speed, at the expense of accuracy. Likewise, editorial decisions are increasingly based on the "attention economy", web analytics, and clickbait headlines rather than on news values and public interest assessment.

This process of decreasing the quality of journalism and the inability of the media to respond adequately to technological change and disinformation challenges has opened space and the need for engaging fact-checkers, i.e., specialised organisations whose task is to verify the accuracy of information.

This study presents an overview of the current state and trends in the global information environment. It specifically conceptualises and elaborates the phenomenon of online disinformation and information disorder, as well as the characteristics, trends and actors of the modern information environment and the relationships between them.

An integral part of the Study is the analysis of the current situation in the field of media, digital platforms, new technologies, journalism, and the functioning of information verification systems. The analysis also includes an overview of the development and potential of the application of computer systems for information verification and fact-checking, as well as the analysis of key international and European standards and criteria in the process of establishing information verification systems, combating disinformation, and strengthening credibility of the media. The Croatian media context is specifically addressed through the dimensions of the media system itself and the media market, trust in the media and other institutions, the characteristics of media audiences and the quality and freedom of journalism in the country.

Based on such a comprehensive analysis, the aim of this Study is to elaborate the basic standards, activities and criteria that should provide information and guidance for designing and implementation of public tenders for activities and projects within the first part of investment C1.1.1.R6-I2 "Establishment of verification of media facts and public data disclosure systems", under component C1.1. 'Resilient, green and digital economy' and reform C1.1.1. R6 "Development of a resilient cultural and creative sector", as part of the National Recovery and Resilience Plan, whose holders are the Ministry of Culture and Media of the Republic of Croatia and Croatia's Agency for Electronic Media.

The National Recovery and Resilience Plan is an action plan which comprises projects, measures, and reforms that EU member states were supposed to develop in order to use part of the resources provided by the special Recovery and Resilience Facility. This mechanism was agreed by the European Commission, the European Parliament, and the European Council in February 2021 to make it easier for member states to exit the economic crisis caused by the coronavirus pandemic, but also to create the foundations for even more resilient societies and more sustainable economies.

Besides the introduction, concluding recommendations and glossary, this study contains the following chapters: (2) Mapping the global environment and defining key terms; (3) Fact-checking organisations; (4) Computational fact-checking systems; (5) Analysis of key documents, international and European standards and criteria and processes for establishing a system for verification of information; (6) Croatia: media, trust and disinformation.

2. Mapping the global environment and defining the key concepts

Disinformation is not a new phenomenon, yet it poses one of the greatest challenges for democracy and informed citizenship in the modern media and information environment.

Informed citizenship is the foundation of healthy democracy. It is achieved by developing media literacy of citizens and ensuring media pluralism. Media pluralism, as the central concept of European media policies, is understood not only through the market dimension and the diversity of content on offer, but also through the distribution of communicative power[1] and an adequate environment (legal and practical) that enables independent, high-quality, and economically sustainable journalism. The rapid development of information and communication technologies and the global impact of online platforms have fundamentally altered the information environment, and, above all, the ways in which news is created, presented, shared, found and consumed.

Significantly lower costs of producing and distributing content online, have broken down some structural barriers for new actors to enter the media markets and, generally, the process of exchanging information that may be of public or general interest. These trends were initially viewed optimistically. They were considered to represent the potential for democratising public debate through the process of opening up public space to individuals or groups who had previously had difficult access or had no access to the media whatsoever. However, recent research points to the risks of information overload (Holton and Chyi, 2012; RIDNR, 2019) and information disorder (Wardle and Derakhshan, 2017), i.e., the situations in which it is difficult to find and recognise credible information.

In order to understand the problem of disinformation and information disruption, it is important to understand the characteristics, trends and actors of the modern information environment and the relationships between different actors.

2.1. Information disorder

DISINFORMATION is not a new phenomenon. However, only in recent years it has been recognised as one of the key challenges for informed citizenship and democracy (European Commission, 2018) and – amid the COVID-19 pandemic – as a threat to public health (European Commission, 2020; WHO, 2020). Disinformation is a complex concept that should be viewed through different dimensions. To be specific, it

1 In relevant literature, 'communicative power' denotes the power to shape or influence public opinion. While such power was previously largely in the hands of media organisations, today it is increasingly in the hands of online platforms and their algorithmic and personalised distribution of media and other content.

needs to be approached by taking into account the dynamics of the emergence of new actors, techniques and technologies of producing and spreading lies and manipulation, especially online. In the key document presenting a framework for European action against disinformation in 2018 - *Communication to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions - Tackling online disinformation: a European approach (COM/2018/236 final)* – the European Commission defines disinformation as verifiably false or misleading information designed, presented and distributed for economic, political or other benefit and with the intention of deceiving the public. As such, disinformation can cause harm and pose a threat to democratic political processes and the public good, such as protecting citizens' health, the environment or security. This definition was adopted by the European Commission on the basis of the High-Level Expert Group on Fake News and Online Disinformation (HLEG)'s proposal. In actual fact, the HLEG is the advisory body, which was convened by the European Commission in January 2018 with the aim of providing the European Commission with advice on the problem of fake news and disinformation spread online and recommending the initiatives needed to combat this phenomenon. In March 2018, the HLEG[2] published its report, in which the complexity of the phenomenon of disinformation was emphasised: (1) although disinformation can be very harmful, it is different from illegal forms of speech, such as defamation, hate speech, incitement to violence, etc.; (2) it is often difficult to separate it from low-quality journalism, clickbait, unintentional mistakes, etc.; (3) both the media and citizens sometimes contribute to its dissemination, albeit without being aware of it; (4) disinformation as a phenomenon goes beyond the notion of "fake news".

Although the very name of the Expert Group contained the term "**FAKE NEWS**", its members subsequently criticised this use and avoided using the term in the final report, citing two reasons:

1. The term is inadequate to encompass the complex problem of disinformation, which includes content that is not actually or completely "false". In reality in the case of disinformation, fabricated information is sometimes mixed with facts, taken out of context, some key information is deliberately omitted, etc. Furthermore, disinformation includes practices that go beyond imitation of "news" and contains various forms of manipulated video materials, memes, infographics, microtargeted advertising, organised trolling, networks of bots, etc.;
2. The term is misleading because it has been appropriated by some politicians and other powerful actors, who use the term to try to discredit the media and journalists who critically report on them. Therefore, frequent and often unfounded use of the term "fake news" by politicians additionally affects negatively the already low trust in the media.

The term "fake news" is used appropriately when it refers to posts and articles based on false information, packaged to look like news, with the aim of deceiving readers for financial, ideological or other gain (Tandoc et al., 2018: 674).

When erroneous, false or misleading information reaches the public space and is spread without bad intention, but because of one's naivety, insufficient knowledge, insufficient verification, speed, fear and other reasons unrelated to deliberately obtaining benefit from it and causing harm, there is a case of misinformation (HLEG, 2018; Wardle and Derakhshan, 2017). The European Commission (2021) also defines it as wrong, false, or misleading information that is disseminated without malicious intent. Nevertheless, the Commission stresses that even in this case the effects can be very harmful. The COVID-19 pandemic has exposed a number of problems caused by the rapid spread of inaccurate, unverified, premature, incomplete, and intentionally manipulated information at various levels.

---

[2] The HLEG consisted of 39 members, led by Prof. Madeleine de Cock Buning. Members included representatives of academia, journalists, media organisations, online platforms and civil society and information verification organisations..

Disinformation and misinformation often overlap in content and differ only in terms of intention, that is, whether someone deliberately disseminates this kind of information in order to cause harm or obtain benefit, or unintentionally, believing that it may be true. Even when incorrect information is spread unconsciously or unintentionally, that is, when citizens, media and other actors who share misinformation believe it is true or do not bother with verification, they contribute to the achievement of socially harmful goals of strategically organised and guided disinformation campaigns. Sharing is not the only way to increase the visibility and reach of disinformation and misinformation. This is also achieved by liking, by commenting, by clicking on certain content, or just by watching posts and content on the screen because platforms measure that as well.

According to the definitions presented above, **THE KEY DETERMINANTS OF DISINFORMATION ARE: VERIFIABILITY, INTENTION AND HARM**. Verifiability, therefore, is one of the primary conditions on the basis of which disinformation can be identified. The principles of this condition can also be found in the practice of the European Court of Human Rights (ECtHR) in Strasbourg, whose task is to ensure that states parties to the European Convention for the Protection of Human Rights and Fundamental Freedoms respect the rights and guarantees provided for in the Convention. While there are no ECtHR rulings that directly address the issue of online disinformation, there are many rulings that address these issues that are applicable to the problem of disinformation as their rationales can be adapted to the online information environment. Article 10 of the Convention guarantees the right to freedom of expression, and this right includes freedom of thought and freedom to receive and transmit information and ideas. Freedom of expression, according to the Convention, is not unlimited and may be subject to restrictions provided for by law. These restrictions are necessary in a democratic society in the interests of national security, territorial integrity or public security, prevention of disorder or crime, protection of health and morals, reputation or rights of others, prevention of dissemination of confidential information or in the interests of preserving the authority and impartiality of the judiciary. Through the Extensive Practice of the ECtHR relating to Article 10, it has been well established that value judgments and personal opinions provide greater protection than verifiable factual claims. The ECtHR pointed out that opinions are less "susceptible to evidence"[3] and therefore cannot be under the same scrutiny as false claims of fact. Proving the veracity of a value judgment would be impossible to fulfil and would call into question the very freedom of opinion[4] . However, it is not always easy to determine the difference between a factual allegation and a value judgment.

Therefore, it is necessary to bear in mind that there is a difference and to consider the circumstances and content of each individual case[5] . In addition, when disputed statements are made during a political debate, elected officials and journalists enjoy a higher level of questioning the actions of the authorities, even when the statements made may not have a clear factual foundation[6] .

Another determinant and identifier of disinformation is the intention of creating and spreading them. In many instances, the main difference between disinformation and misinformation does not lie in the very content being shared, yet it arises from the fact that there is intention and awareness of the spread of false or misleading information. However, even when there is no intention, spreading erroneous or manipulated disclosures can result in harm, both to individuals and society. Online platforms and digital communication have amplified the potential for harm due to their reach and information flow speed.

---

3 ESLJP, McVicar v. the United Kingdom: 46311/99; (2002);, ECtHR, Lingens v. Austria: 9815/82; (1986)

4 ESLJP, Morice v. France: 29369/10; (2015) [GC]; ECtHR, Dalban v. Romania: 28114/95; (1999) [GC]; ECtHR, Lingens v. Austria, 9815/82; (1986); ESLJP, Oberschlick v. Austria (1): 15974/90; (1995)

5 ESLJP, Brasilier v. France: 71343/01; (2006); ECtHR, Ballaskas v. Greece: 73087/17; (2020)

6 ESLJP, Lombardo and Others v. Malta: 7333/06; (2007); ECtHR, Dyuldin and Kislov v. Russia: 25968/02; (2007)

Legal doctrine related to, e.g., defamation always considers whether a particular opinion is expressed publicly or privately. Therefore, public exposure of expressed opinions and the scope of effect of the platform define the potential punishment.[7] The European Court of Human Rights, on the one hand, stresses the potential of the online environment for exercising freedom of expression. On the other hand, it warns that the risk of harm is also growing significantly, precisely because of the incomparably greater reach of online platforms compared to traditional media, especially in terms of the right to respect for private life and reputation[8].

Although there is fundamentally agreement at the European level on the basic definition of disinformation, the problem arises with operationalisation as it is an extremely complex phenomenon that encompasses different dimensions.

Experts and scientists generally distinguish at least four dimensions of disinformation that require attention: (1) harmful content, (2) manipulative actors, (3) the diversity of techniques and technologies used by actors to increase the reach and impact of their campaigns, and (4) the digital architecture of online platforms and their internal policies that affect the potential spread of disinformation.

A study conducted by Claire Wardle and Hossein Derakhshan for the Council of Europe in 2017 laid the foundations for understanding the complexity of the information disorder we are in, as well as for studying the short- and long-term consequences that such a disruption can have on political and other social processes, deliberative communication and trust in different institutions. Their conceptual framework looks at information disorder through:

- three types of information disorders: disinformation, misinformation (erroneous information) and malinformation (malicious information based on reality but used primarily to harm a person, organisation or state; often implying the intentional disclosure of private information which is not in the public interest, such as revenge pornography);
- three stages of information disorder their creation, media production and distribution;
- three elements of information disorder: the creator (the one who creates and shares it), the message and the interpreter (recipient).

The creators of disinformation campaigns are various, just like their motivation and effectiveness. Thanks to financial and other resources, political actors, notably the ruling ones, have the greatest potential to create influential misleading campaigns. Since Russia's invasion of Ukraine in 2022, but also earlier (Brexit and the 2016 U.S. presidential election), much has been written and talked about in terms of the impact of Russian propaganda on information environments and democratic processes in Europe, in the United States and beyond. In 2015, the European Union set up the East Stratcomm Taskforce - a task force that regularly analyses and reports on disinformation coming from Russia. Stratcomm Taskforce reports show that the key strategy of the authorities in Russia is to spread as many conflicting messages as possible to confuse the audience and convince them that there are too many versions of events and that there is not just one truth. Research continuously finds evidence of the state's use of disinformation to achieve certain domestic-political or foreign-political goals. King, Pan and Roberts (2017) documented that China paid people to post millions of fictional

---

7 Media Legal Defence Initiative and International Press Institute, 'Freedom of Expression, Media Law and Defamation' [2015] <**https://www.mediadefence.org/wp-content/uploads/2020/06/MLDI.IPI-defamation-manual.English-1.pdf**

8 ESLJP, Editorial Board of PravoyeDelo and Shtekel v. Ukraine: 33014/05; (2011)

social media posts with a view to diverting public attention from political issues that could fuel discontent and protests. Likewise, a survey by The New York Times and ProPublica, a non-profit newsroom investigative journalism[9] revealed how Chinese authorities organised an army of "trolls" (fake Internet commentators) in an attempt to control the coronavirus-related narrative on social media as early as in January 2020, that is, at the beginning of the coronavirus pandemic. They wanted to create the impression that the virus looked less serious and that the authorities were capable of dealing with it (Zhong et al., 2020). Research by Zhong et al. (2020) revealed that some of the tens of thousands of fake commentators earned $25 for original posts, 40 cents for flagging negative comments for deletion, and one cent each for sharing desirable posts. However, Russia and China are not the only countries participating in this kind of disinformation campaign. A study by the University of Oxford, published in 2019, presented evidence of the existence of organised campaigns to manipulate public opinion through online platforms in 70 countries of the world, and they also showed that in each country there is at least one political party or government agency that uses social media with a view to influencing public opinion (Bradshaw and Howard, 2019).

Disinformation amounts rarely to a brand-new item of lie and fabrication as it is more demanding and expensive to create it as such (Wardle and Derakhshan, 2017). Much more often it is about re-elaborating pre-existing content that is put into a false relationship (e.g., a title that is not a reflection of the text at all), false context (using photos from one event to illustrate a completely different one, without clearly specifying it), or other types of intervention in the original content (e.g., using photoshop or deepfake technology in order to tamper with the original audio-visual materials, e.g., by replacing the person's face with someone else's in an illustration or intervening in speech so that it turns out that somebody said something they did not originally say). Visuals (such as photographs, illustrations, infographics, memes) can be far more convincing than other forms of communication (Birdsell and Groarke, 1996), which makes them more effective forms to be used for the purpose of disinformation and misinformation (Kietzmann et al., 2020; Wardle and Derakhshan, 2017).

Manipulations manifest themselves differently on different platforms, depending on the specific architecture and concept of the platform itself. For example, in order to be effective on TikTok, they must be in the format of attractive and short videos. On Facebook, on the other hand, better reach is achieved by paid disinformation campaigns (ads) (Chiou and Tucker, 2018), while on Twitter they are very successfully spread even organically (Vosoughi et al., 2018). In closed messaging platforms and apps (like WhatsApp, Viber) people fall for erroneous and manipulative posts even more easily as they get them directly from people they know.

The extent to which disinformation really is a problem depends on a specific national context that includes, among other things, a broader legal tradition, a level of media literacy, a political context, a level of trust in institutions, and the quality and status of traditional, especially public, media. What all these harmful content items and patterns of behaviour have in common is that they create an information disorder in which it is difficult for citizens to distinguish credible from misleading information. Such an environment encourages distrust in traditional media and other institutions, and positions alternative sources of questionable credibility as the main "sources of truth" for a significant number of social media users (Nenadić and Vučković, 2021).

In such an environment, the system of filtering, organising, and presenting relevant and credible information seems more important than ever before.

---

9 See: **https://www.propublica.org/article/leaked-documents-show-how-chinas-army-of-paid-internet-trolls-helped-censor-the-coronavirus**

2.2. Media and journalism - part of the problem and part of the solution

In an environment of information disorder, the media are simultaneously part of the problem and victims of such an environment. For most of the twentieth century, the business models and professional routines of journalism in developed countries were very stable (Lewis, 2012). The media were central institutions in the process of informing the public, which resulted in their privileged status as the fourth estate, their social role and their responsibility. The development of information and communication technology, and especially the growth of social media, has changed the information environment. The media are no longer the only institutions of public reporting and editorial mediation. Although their role remains important, several new actors and processes affect journalistic practices, the distribution of news and media content, their consumption and the effectiveness of business models and the economic sustainability of the media. Even basic definitions are being questioned, that is, who a journalist is and how to define the media.

Smartphones and online platforms have enabled users to generate and share a wide variety of content. Although not all the content that is shared online has a journalistic value or a news value, the fact is that almost anyone can report on almost anything with potentially global reach. Online platforms make it easier for various individuals and organisations to actively contribute to the processes of informing the public and news exchange.

This significantly accelerated an already rapid information process that leaves little time to verify information before publication. It poses a significant challenge to professional journalism which, in order to be current, has an imperative of speed, sometimes at the expense of accuracy. Social networks and the information circulating through them have become an indispensable source of journalistic reporting, especially in emergency situations, without developing clear practices of professional activity of journalists in such an environment, especially in the context of verifying information and sources (Nenadić, 2020a).

It should be borne in mind that online platforms, such as social networks, search engines, news aggregators and messaging applications, in most cases neither legally nor professionally have editorial responsibility yet and anyone can share almost anything on them with a potentially wide reach. In addition, different platforms have different policies concerning moderation and ranking of content and define differently what is acceptable behaviour and what is not allowed.

Thus, on the one hand, these digital platforms can be valuable sources of information, especially in situations where journalists find it difficult or impossible to access the location, and at the same time, due to their architecture and the nature of business and the way in which different actors use them, they are effective channels of spreading disinformation.

Since it is no longer adequate to define journalists solely through their affiliation with a media organisation or through their membership of a professional association, many experts highlight professional standards and ethical principles as the main difference between journalists and others who act in a similar way. In the digital age, one-time journalism can be produced by anyone. Professional journalism, on the other hand, is viewed as a more systematic process based on accumulated and applied practical knowledge. Its value derives from its goals, practice, and ethics (Davis, 2010). Accordingly, commitment to the public interest, supervisory function, objective methods and verification are commonly used to underpin the value of journalism and justify the profession's authority in society. The requirement to verify information is highlighted in particular in the context of digital platforms and abundance of information. The American Institute of Media, an educational non-profit organisation with a long history, describes journalism as a "discipline of

verification" - fact-checking and evaluating the credibility of sources. This is what makes journalism different from other forms of communication and distribution of information[10].

In the traditional media system, news was an integral part of products (newspapers, magazines, television, or radio newscasts) and editorial processes that, at least normatively if not always practically, guaranteed compliance with certain professional standards. In an online environment and through social media distribution, news becomes disconnected from the media and "lives a life of its own". This separates to some extent news from established journalistic and editorial processes and allows other actors, without editorial and professional responsibility, to disseminate information, but also disinformation, and shape narratives. The process of distribution and fragmentation of communicative power in an online environment is impossible to stop, and it would be undesirable from the perspective of freedom of expression. Yet, it sets the imperative to redefine the role of journalism and media organisations from which professional selection and presentation of important and credible news is increasingly expected. Journalism has always been a special profession in which open access to the profession is one of the fundamental standards and principles of protecting media freedom: formal journalism education is not a requirement for engaging in journalism, whereas imposing any restrictions on journalistic practice, such as licensing journalists, would conflict with democratic constitutional theory that fosters freedom of expression as a fundamental right (Wenger, 2006).

Media systems operate within broader social systems that cross national borders especially in a globalised world. The rapid diffusion of new communication technologies calls into question power relations, allowing more direct mass communication (many towards many) and denying traditional media organisations exclusivity in terms of deciding what to report as news. Technology companies acting as digital intermediaries (e.g., social networks, news aggregators and search engines) are becoming dominant actors in the distribution and process of finding and consuming news (RIDNR, 2017), and completely dominate digital advertising, which is still the main source of revenue for the media in the digital environment as well.

Media companies have traditionally acted as a two-sided market, combining revenue from advertisers and audiences (readers, viewers, listeners). However, this business model has not successfully transferred to a modern online environment in which digital intermediaries (platforms) dominate online advertising. Thanks to global popularity and the collection of detailed user data, platforms enable more effective advertising that targets segmented audiences and has a wider reach. In doing so, advertising funds that once served to fund journalism have been significantly diverted to online platforms. The media are trying to offset losses and prevent advertising outflows by offering advertisers, among other things, so-called native advertising and promotional services that are opposed to the traditional principle of clearly separating news from ads (see Carlson, 2014; Küng, 2015; Laursen and Stone, 2016; Matteo and Zotto, 2015).

At the same time, a "culture of free" was being built on the Internet in which people expected open access to various content (see e.g., Goyanes 2014; Hill and Lashmar 2013). However, in the last few years it has become increasingly clear to both the media and the audiences that quality journalism has high production costs and that the media, in order to ensure economic sustainability, must increase the willingness of audiences to pay for the content they consume online. A Reuters survey of digital media trends shows that people's willingness to pay for online media content is growing in some countries (RIDNR, 2022), especially when people feel that the media are providing them with credible and timely information that allows them to understand the world in which they live. This is the case, for example, in the Scandinavian countries (according to the last RIDNR from 2022, especially in Norway and Sweden),

10 See: **www.americanpressinstitute.org/journalism-essentials/verification-accuracy/journalism-discipline-verification/**

which can probably be attributed to a higher standard of living, but also to the highly developed media and political culture of these societies. However, the Reuters survey also shows that citizens, for economic and other reasons, generally pay for a subscription or access to only one or a few of the leading media brands (RIDNR, 2022). In addition, audience surveys also suggest that citizens are more willing to allocate funds for innovative media platforms and entertainment content such as those provided by, for example, Netflix and Spotify, than for serious media and journalism (Kantar Media for Reuters, 2017).

In addition to subscriptions or one-time donations to access media content, the media also experiment with paywall systems which make certain content available only after payment. There are different types of paywall: from billing for all content on a particular website, for example on a portal, charging only for special content and forms that require significant work and expert analysis, a measurement model that offers free access to a limited number of articles over a given period of time, and everything after is charged, to micropayments for access to individual articles within one publication or from various publications, by their own choice (Casero-Ripollés and Izquierdo-Castillo, 2013; Nenadić and Ostling, 2018; Picard, 2014).

The economic instability of the media is associated with the economic instability and precariat of journalists. In such a context in which journalists lose their jobs, are paid less and work in unstable arrangements, while newsrooms rely significantly on lower-paid, less experienced journalists, journalism loses quality, editorial autonomy, and the potential to cope with the never greater challenges of verifying information and publishing credible information (see, e.g., Cohen, Hunter and O'Donnell, 2019; Compton and Benedetti, 2010).

However, the media themselves are not exempt from responsibility for their current state of affairs. Due to short-sightedness and short-term profits, commercialisation and sensationalism, insufficient investment in research and development, they failed to respond to the opportunities and challenges of emerging technologies and new habits of audiences (see, e.g., Compton and Benedetti, 2010). In addition, scientific research increasingly warns that the mainstream media play a significant role in spreading inaccurate and misleading information online (Allen et al., 2020; Tsfati et al., 2020). This is associated with the decline of media and journalistic standards and practices. On the other hand, it can be attributed to a special status and greater visibility that content generated by traditional (or mainstream) media has when published on social networks (due to social status and algorithmic policies of organising and ranking content).

Mistakes and manipulations can get into the mainstream media due to the lack of verification, clickbait strategy, i.e., the imperative of "attention economy" (Lischka and Garz, 2021) or the political dependence of the media and the bias of their editorial policies. Even when the media report on disinformation in order to expose it, if they do so in an imprudent way, they can contribute to its spread, visibility and legitimacy. In reality, a part of the audience, due to selective attention and cognitive biases, can still retain only wrong information, which has been confirmed by research.

For example, as Tsfati et al. (2020) show in their review and synthesis of literature exploring the role of the mainstream media in disseminating inaccurate and misleading information, the media actually play a significant role in the dissemination of misinformation and disinformation. In addition to unintentional mistakes, clickbait headlines and poor journalism, circumstantial evidence presented by Tsfati et al. (2020) suggests that more people learn about fake news from the mainstream media than from social media. This is in line with the "Trumpet of Amplification" (Wardle, 2018), which shows in a graphic way that manipulations mainly start from smaller marginal groups that first share them in their closed groups or in conspiracy theory supporters' communities.

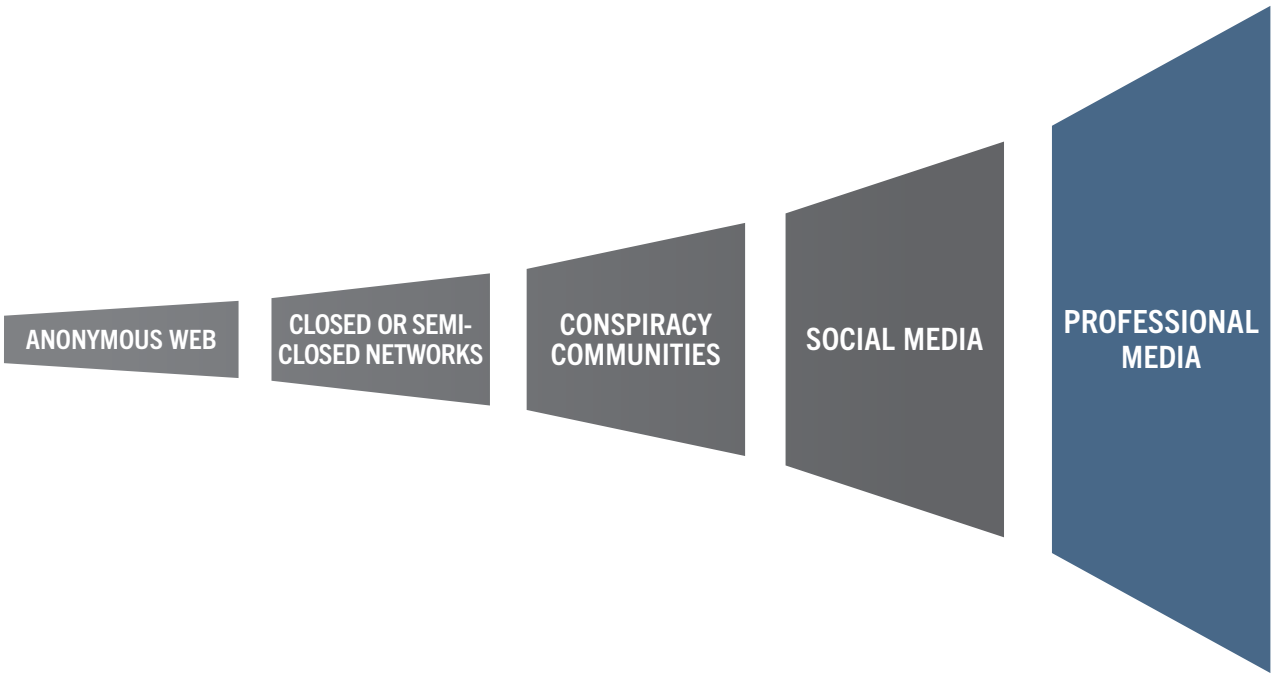


Figure 1: The Trumpet of Amplification. **Source:** Wardle 2018

The first big success that disinformation and its creators achieve is when they start spreading through social networks, and especially if it successfully moves from platform to platform. However, the main goal is achieved if disinformation is taken over and published by traditional media that thereby give it oxygen, legitimacy and present it to a wide population. It is very difficult to refute untruths and manipulations after they have been published by the traditional and mainstream media.

In the context of information disruption and relativisation of science and scientific evidence, the golden journalistic standard of balanced reporting, i.e., presentation of different sides in the topic being addressed, has also come under fire. Misinterpretation of this standard has resulted in a false balance, i.e., giving equal space to opposing views even though they are not at the same level of expertise, nor are they equally supported by scientific evidence (Imundo and Rapp, 2021). A paradigmatic example of such an approach is climate change reporting that is increasingly framed as conflict between "warners" and "deniers", who are treated as two equal sides in the discussion. Consequently, important and substantive discussions and topics on climate change are lost from focus (Brüggemann and Engesser, 2017).

Creators of disinformation campaigns themselves sometimes use mainstream media, their clickbait headlines, and reporting based on speculation rather than on expert analysis to validate or strengthen their theses (Soares and Recuero, 2021).

What poses one of the key challenges for professional journalism and media today is the question as to how to bridge the gap between the current state of affairs and the expected role of journalism in

society, especially if we take into account the fact that, to some extent, the media as the key fact-checkers have been converted themselves to objects of verification by specialised organisations and fact-checking programmes. Trust in traditional media in several countries, including Croatia, is in a significant decline.

The latest 2022 Reuters Institute survey shows that average trust in news at the level of all 46 countries included in the survey is 42%, down from 44% in 2021. The highest trust in news was recorded in Finland (69%), and the lowest in the United States (26%). Croatia is among the countries that recorded the biggest drop in trust in news according to Reuters measurement - from 45% in 2021 to 38% in 2022. Likewise, RIDNR has seen a significant average decline in news interest – from 63% in 2021 to 51% in 2022. Such a decline can be explained in part by the greater interest in news related to the 2021 pandemic and the audience's fatigue by that same type of news and the topic in 2022. However, it is reasonable to believe that this is a trend which is significantly influenced by factors related to changes in the media and in the habits of audiences. This is supported by the fact that the number of people who actively avoid the news has grown strongly. This percentage, according to RIDNR (2022: 13), averaged 29% in 2017, while in 2022 it was 38%. As the most important reasons for avoiding news, respondents cite monotonous topics in news broadcasts, i.e., too much focus on politics and COVID-19 (43%), then negativity of news, which badly affects their mood (36%), while 29% of respondents claim to avoid news because they are biased and cannot be trusted.

While trust in traditional media is declining, it is still greater than trust in the news and information that people find on social media (Eurobarometer 95, Spring 2021; Eurobarometer 96, winter 2021/2022; RIDNR, 2021, 2022).

This leaves room for the media and journalists to increase trust, not solely competing over who will be first, but by distinguishing themselves from other information providers by offering complete, verified, accurate and credible information.

2.3. Digital media and online newsrooms

The past two decades have witnessed a paradox characterised, on the one hand, by the exponential development of media devices, platforms, and services, and on the other hand by the increasing degradation and instability of economic conditions for journalists and other media workers (Deuze et al., 2010). The Internet and new communication technologies have enabled and accelerated the advancement of traditional media into the online environment, as well as the development of original digital media, journalistic projects and various platforms for aggregation and distribution of news and other content. At the same time, a report on the state of technology in newsrooms around the world, published in 2019 by the International Center for Journalists[11], shows that more and more journalists are losing their jobs due to reorganising newsrooms or cutting costs or are forced to work as freelance journalists in volatile engagements.

The growth and development of online news and media delivery have broken down time and spatial barriers and created an environment where citizens expect news to be delivered at all times, no matter where they are. Media organisations adapted to these trends to include and develop new content distribution platforms (e.g., daily newspapers developed online editions, and then their own

11 See: **https://www.icfj.org/sites/default/files/2019-10/2019%20Final%20Report.pdf**

audio-visual channels). At the level of organisation of newsrooms and production of media content, this resulted in convergence, that is, merging production for different platforms into one newsroom. In doing so, media organisations expected journalists to possess and further develop skills in preparing content for different platforms without providing them with adequate and continuous education and workshops (see, e.g., Singer, 2004).

Accelerated and continuous news production cycles have also been reflected in the quality of media content, especially since online newsrooms often collaborate with journalists with less experience and with rather unstable contracts. A series of studies have shown that, especially with digital media, journalists are rarely leaving the newsroom to report from the field, and are increasingly recycling, i.e., transmitting various press releases and translating texts from foreign media – the so-called "churnalism" (e.g., Davies, 2008; Kovach and Rosenstiehl, 2001; Lewis et al. 2008). This situation in converged and online newsrooms has raised concerns for decades about the "deprofessionalisation" of journalism (Bromley, 1997).

In recent years, journalism's autonomy and professionalism have been further jeopardised by economic pressures, according to which editorial decisions are based more on web analytics and clickbait headlines than on the value of news and public interest assessment (see Blom and Hansen 2015; Lischka and Garz, 2021; Tandoc 2014).

The accuracy achieved by checking the source and the content itself is one of the key journalistic standards. On the other hand, speed is one of the main characteristics of the online information environment. With the growing impact of social networks and the acceleration of the news cycle, especially in digital media, journalistic standards of accuracy and verification of information have been degraded (Hermida, 2012). When a news story appears on one of the online platforms, journalists cannot stop its spread and do not have time to check it out. This puts journalists in front of a dilemma: being first or being accurate. While normative standards unanimously require verification of information before publication, the practice is often different. The tension between speed and accuracy in journalism is not new, but it has been intensified by the emergence and relevance of the Internet and social media. Kovach and Rosentiel (1999) argued that even before the advent of the leading social media, the incessant news cycle made journalism more inclined to publish allegations and speculation, rather than establishing factuality. There are several examples in which the world's leading media organisations failed in accuracy while competing in speed (Nenadić, 2020a). It is clear that speed is one of the imperatives of media organisations if they want to maintain relevance in the news and information market. However, it is becoming progressively clearer that information verification has even greater importance in an environment characterised by an abundance of information and concerns about the extent and impact of disinformation. Among the skills expected from journalists and media organisations today, there is the use of digital information verification tools and the development of automation and artificial intelligence-based technologies that are able to speed up the verification process. However, a survey by the International Centre for Journalists (2019) found that although journalists are increasingly using digital tools to verify information[12], these tools are still not widely used. In addition, training and workshops for journalists in this area are stagnant or even decreasingly available, and the use of artificial intelligence in the verification process has not reached the level of simple and reliable application.

12 E.g., tools for identifying reliable news sources, reverse image search to verify the source of photos and other illustrations, a platform for checking content from social networks, and the like.

Journalists are increasingly using social media for a variety of purposes, including collecting and sharing information, networking, but also for promoting their work or the organisation they work for. In addition, they often do not separate the private from the business use of social networks, which, on the one hand, contributes to authenticity and transparency, while on the other, it potentially negatively affects the public's perception of journalistic impartiality (Nenadić, 2020a). Although journalism has a long tradition of self-regulation, in most EU countries, there are no general and publicly available codes or guidelines that would recommend or prescribe how journalists should use social networks (Brogi et al., 2020). If they exist, provided they are publicly available, such guidelines are often criticised for restriction of journalistic freedom of expression or because of the process by which they were created, and which is mainly managed by the marketing departments of media companies, with the limited participation of journalists themselves, which raises the issue of editorial autonomy (Nenadić, 2020a). The latest Reuters survey (RIDNR, 2022: 18) shows that more than half of respondents in the vast majority of countries covered by the survey believe that journalists on social media should write about news rather than express their own opinions. However, RIDNR (2022) points to a generational discrepancy and states that younger audiences are much more inclined to believe that journalists on social networks have the right to express their own opinions, than is the case of the audiences over 45.

2.4. Technology companies as digital intermediaries

The media remain key institutions in the production of news and other content of general or public interest, but online platforms, mobile applications and smart devices have taken on the role of media intermediaries between media and their audiences. With their design and algorithmic sorting and ranking of content, these digital intermediaries shape an information environment in which citizens interact with news from both media and other sources.

Online platforms, such as social media[13], search engines[14], news aggregators[15] and messaging apps[16], are the primary sources of information for an increasing number of people, especially young persons who grew up with social networks (so-called social media natives) (RIDNR, 2022; Standard Eurobarometer 96, 2021/2022). Platforms generally do not produce their own content, but distribute content produced by their users and algorithmically create personalised recommendations of media, political and other content, in accordance with the characteristics and preferences of individual users. Platforms, therefore, act as key intermediaries in the process of distributing and consuming news and other content produced and shared by their users. There are various social media users who use platforms in vastly different ways and with different goals. For example, Facebook users are both ordinary citizens and media and politicians and political parties and marketing agencies and many others. Users of online platforms are also journalists and media organisations who, on the one hand, use them as a source of information, and, on the other, for the distribution and promotion of their work. Online platforms are primarily private, mostly American corporations, driven by neoliberal market values, but at the same time they are also key actors of the public information sphere and democratic systems around the world (Van Dijck, 2020). They, unilaterally and without public oversight, create an architecture in which users find and interact with news (Helberger, 2016).

13 E.g., Facebook, Instagram, YouTube, TikTok

14 E.g., Google, as the leading search engine

15 E.g., Google News

16 E.g., WhatsApp, Telegram, Signal, Messenger

Given the great popularity and global user base, social networks and other platforms have the ability to focus the attention of a large number of users on certain issues and push other topics into oblivion. In this way, online platforms, among which Facebook remains the leading one, shape the terms of public debate.

Some authors, such as Natali Helberger (2016) describe this as a "privately controlled public sphere", demonstrating a new form of vertical integration in which gathering a large number of user data and their use play a central role (Nenadić, 2020b).

Algorithmic content recommenders are considered necessary to enable users to navigate the vast amount of information available online (Gauch et al., 2007; Helberger, 2009; Oulasvirta and Blom, 2007). However, a growing number of authors also see them as a potential threat to pluralism and democracy, because their actions are not transparent (Zuiderveen Borgesius et al., 2016), and because personalised content selection may fragment the public sphere, thus preventing users from being exposed to different views and opinions, and denying the opportunity for democratic dialogue (Pariser, 2011; Sunstein, 2002, 2009).

In addition to recommending content, platforms moderate it according to certain criteria (filtering and removing) and rank it (giving greater visibility to certain content or sources). One of the criteria in this process is to meet the tastes or expectations of individual users because their stay on the platform and greater engagement also increase advertising revenue (Gillespie, 2018).

The business model of online platforms is therefore based on the greater engagement of users and the retention of their attention, which is then sold to advertisers. Such a business model gives greater visibility to content that achieves greater engagement and has the potential of virality, and as shown by Vosoughi et al. (2018), it is falsehood, especially political falsehood, that diffused significantly farther, faster, deeper, and more broadly than the truth.

Apart from the fact that the business model of online platforms favours the dissemination of problematic, provocative and harmful content, such as disinformation, platforms with their policies and business decisions shape the information environment and conduct a significant part of the online distribution of media content. When their dominant position in the digital advertising market is taken into consideration as well, it is clear that the internal decisions and practices of the platforms directly affect the economic viability of the media, and indirectly – by mediating between the media and their audiences – informed citizenship. Due to the current lack of legislative framework[17] and regulatory oversight of the operation of platforms, it is left to them through self-regulation to define and treat the problem of disinformation, but also to decide which sources are credible, and thus more visible or higher ranked in the organisation and presentation of content to individual users.

Some platforms, such as Snapchat, and in some countries Facebook (where it has special news feeds) separate news from other content, while, for example, on Instagram and TikTok, news merges with other videos and photos that users share.

---

17 The Digital Services Act (DSA) entered into force at the time of conclusion of this study. The DSA is expected to take full effect in early 2024.

2.5. Audiences

Not only is news created in a different way, but it is also found in other places and through a different system of presentation and ranking. As the regular Reuters Digital Trends Report (RIDNR, 2021) has shown for years, citizens who gather information online are increasingly doing so through intermediaries - online platforms, and less and less directly from the online media. In 2017, about 65 percent of respondents (or 73 percent of under-35s) from the 37 countries surveyed preferred to access news through search engines, social networks, news aggregators, emails, or mobile notifications rather than directly through the media (RIDNR, 2017).

Research has also shown that citizens do not use online platforms primarily for information, but for other purposes (e.g., for entertainment). However, during the time they spend on platforms, they are exposed to different information and news that potentially shape their view of reality and opinion (RIDNR, 2021).

The Reuters report (2021) shows a stable and strong position of social media as a news source, particularly among younger people and those with lower levels of education. Interestingly, young people under 35 on platforms like TikTok, Snapchat and Instagram as key sources of 'news' do not identify journalists or even politicians but "Internet personalities", known as influencers.

A Reuters survey from 2022 shows that Facebook usage is slowly declining, while Instagram, TikTok and Snapchat are seeing strong growth (RIDNR, 2022: 24). Such a decline in the use of Facebook can be primarily explained by the behaviour of young audiences who belong to the group of social media natives, who prefer visual platforms intended primarily for the exchange of photos and video content. Judging by the 12 countries analysed in more detail by the Reuters survey, Facebook is still the platform that respondents most often use for news, although that percentage is also declining. Twitter is more often perceived and used as the primary destination for news, while YouTube, Instagram, Snapchat and TikTok are primarily used for entertainment, but sometimes they also serve as a source of news and different perspectives (RIDNR, 2021). In recent years, their role in (visual) mediating debate on serious topics such as mental health, climate change, COVID-19 and racial equality has also been noted (RIDNR, 2021).

The Reuters survey from 2022 draws particular attention to the growing influence of TikTok, especially among the population of social natives, but not only among them (24-25). One of TikTok's strengths over other platforms, according to the report, is in an algorithm that selects content based on what the user likes and what others watch, and less on the basis of who the user follows.

The Reuters survey from 2021 shows that when people use social media for news on Twitter and Facebook, they pay most attention to mainstream news, on YouTube mainstream news and personalities (influencers), while on Instagram, TikTok and Snapchat they pay most attention, in the context of news, to personalities (stars, influencers etc.).

Online platforms are full of information and opinions shared by ordinary people, advertisers, activists, politicians, media and other profiles that the user follows, or the platform recommends. According to the Reuters survey (RIDNR, 2021), Facebook and Twitter users from the U.S. pay most attention to news coming from mainstream media and journalists. Furthermore, politicians and political activists, who often use online platforms to bypass the mainstream media are a significant source of news to users of platforms like Twitter. In general, the findings of the Reuters survey suggest that people who trust traditional media less are more likely to seek and pay attention to alternative sources, especially on YouTube.

3.
Fact-checking organisations

As it has been pointed out in the foregoing chapters, with the advent of the Internet, and especially with the rise of social networks, the processes of production, dissemination and consumption of news, information and entertainment have changed. The role of the media as "gatekeepers" that decide what information and in what format will reach the audience has been considerably reduced or at least transformed. The Internet and social networks have made it possible for anyone with access to the network to create and share their own content. Thus, audiences, i.e., citizens are significantly empowered in terms of participation in communication processes. However, at the same time, this democratisation of the information space has led to an explosion of disinformation, as well as to the emergence of strategic manipulation of content with the aim of pursuing certain particular interests and obstructing democratic processes. The Brexit campaign that advocated the withdrawal of the United Kingdom from the European Union and the 2016 U.S. presidential election represent paradigmatic – and at the same time the most famous – examples of such strategic manipulation of information and production of "fake news" (Bennett and Livingstone, 2018). The systematic dissemination of fabricated information and conspiracy theories that garnered huge followings on the Internet, such as QAnon, significantly affected political polarisation after the 2016 election in the United States. It also contributed to one of the biggest incidents in American political history – the Attack on Capitol Hill in January 2021, when supporters of Donald Trump flocked to the U.S. Congress, trying to obstruct the process of counting electoral votes, whose purpose was to confirm the Democratic candidate Joe Biden's victory in the 2020 presidential election.

These events and trends in the media environment in general, encouraged by the development of digital technology and the rise of online platforms, had built the awareness of the need for a systematic approach to the verification of information in the digital space even before the outbreak of the pandemic. These events also indicated the necessity for the regulation of platforms on which it is possible to publish fake content and create disinformation networks that, in their ultimate consequence, seriously threaten democratic processes. With the advent of the pandemic in 2020, this "disinformation order" (Bennett and Livingstone, 2018) became an acute social problem, and with the outbreak of war in Ukraine in 2022, the problem of disinformation assumed a new dimension – from the use of disinformation for the purpose of malicious propaganda activities, to the inadvertent (and irresponsible) publication of unverified and inaccurate information on social networks and in the media.

One way to combat the "bogey of disinformation" has been the introduction of a large number of independent fact-checkers in many states.

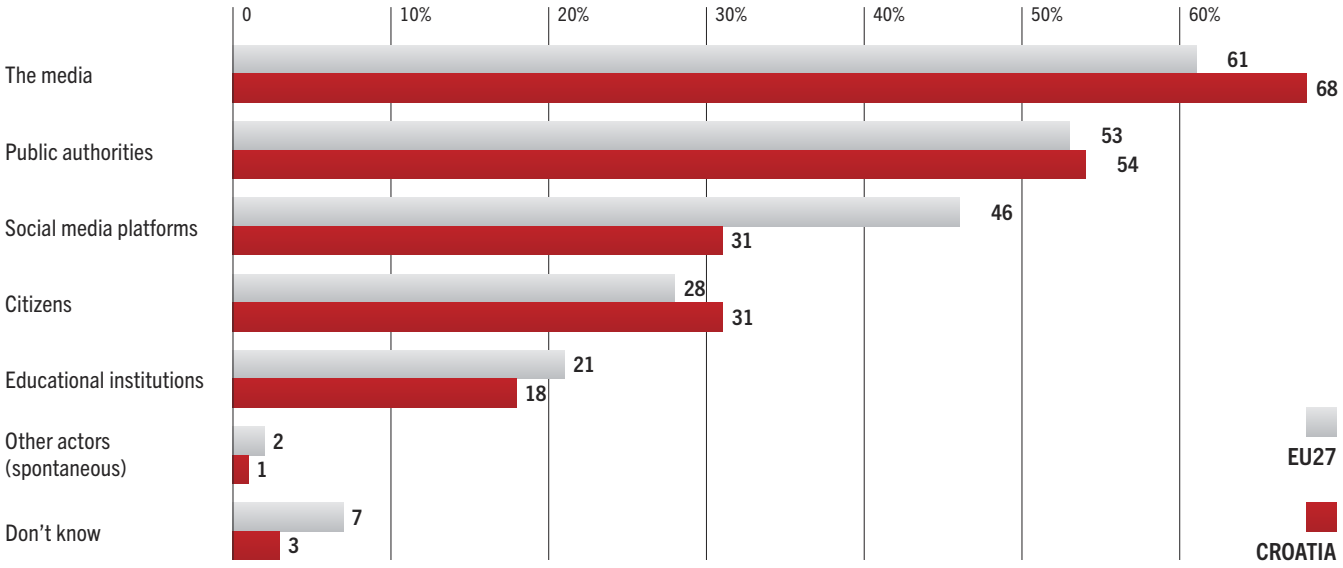
The rise of fact-checking organisations is generally attributed to the watershed moment in the evolution of democratic societies which is characterised by three processes: the decline in the quality of journalism, technological change and socio-political crises and/or reforms (Amazeen, 2020, 2018).

In this sense, Lucas Graves (2016), an author who deals extensively with the phenomenon of fact-checking, attributes the rise of the fact-checking movement in America primarily to the breakdown of "traditional objective reporting" and believes that three key factors contributed to the development of fact-checkers: changes in journalistic standards and practices, technological transformations that have largely robbed traditional media of the role of "gatekeeper" and ultimately, the process of narrowing the space for meaningful public debate in a shattered and fragmented media environment.

"Specialised" fact-checking is not a new phenomenon. The precursors of today's fact-checking organisations were independent initiatives created in the United States some twenty years ago that aimed to check claims and content appearing in the American political space, e.g., in election campaigns. One of the first such organisations is the FactCheck.org founded in 2003, whose aim was and remains to be monitoring "the factual accuracy of what is said by major U.S. political players in the form of TV ads, debates, speeches, interviews and news releases."[18] Their goal, as they claim, is to apply best practices in journalism and science to increase public knowledge and understanding of political processes.

Although determining accuracy, i.e., factuality, is inseparable from journalistic work (or at least it should be), there are some essential differences between journalism and fact-checking. First of all, journalists, i.e., the media, should check the information before publishing it. Therefore, one of the assumptions of good journalism is that it is factually true, while fact-checkers publish claims that they have found to be incorrect and subsequently deconstruct or debunk them. It is interesting, however, that citizens, according to some research, believe that the exponents of the *fight against disinformation* should primarily be – the media. According to the Eurobarometer survey conducted at the end of 2019, this is the opinion of 61% of respondents across the European Union and 68% of respondents from Croatia (see Graph 1).

In your opinion, which of the following should be responsible for combating fake news or disinformation?
(multiple answers possible)



Graph 1: Perception of the bearers of the fight against disinformation. **Source:** Special Eurobarometer 503, December 2019

18 See: **https://www.factcheck.org/about/our-mission/**

Such an understanding can probably be linked to the perception of the media as the main actors of informative activities, who should, by definition, verify information and establish facts before reporting on it. Likewise, a number of media, in addition to regular editorial, professional routines of verifying information have separate fact-checking projects (see, e.g., Fact Checker of the Washington Post[19]).

As the number of specialised organisations and projects for determining the factuality of posts and statements grows, the complexity of the relationship of these organisations with traditional media is becoming more visible, especially since fact-checkers sometimes put under scrutiny the work done by journalists. As online platforms increasingly hire fact-checkers to try to reduce the amount of disinformation and misinformation on their platforms, this sometimes implies the removal of certain content from traditional media sources, which fact-checkers find to be not factual or, e.g., that the headline does not match the text and can be manipulative. Moreover, UNESCO[20] pointed to such a complex and problematic relationship between the media and fact-checkers, the actors that are supposed to be partners. In 2021, UNESCO focused its activity of building trust in the media on the issue of strengthening the relations between fact-checking organisations and traditional media, notably in Southeast Europe. In this area, UNESCO marked occasional complaints by the media about the manner in which the fact-checking process and evaluation of media content circulating on online platforms is carried out. Specifically, fact-checkers, to some extent, took on the role of monitoring the work of the media.

The project of building trust in the media in Southeast Europe and Turkey, jointly launched by UNESCO and the European Union, in which the European Journalists' Association participates as well, is an attempt to respond to the rapid digital transformation and information chaos marked by the degradation of professional journalistic standards, the proliferation of disinformation and a clear decline in civil society's trust in the media. Although the project does not cover Croatia, it is implemented in neighbouring countries such as Bosnia and Herzegovina, Serbia, and Montenegro, and also includes Albania, North Macedonia, Kosovo and Turkey. The aim of the project is to develop media and information literacy skills in formal and non-formal education, to improve the capacity of the media to combat disinformation and misinformation, and to strengthen media accountability and sustainable and effective self-regulation mechanisms[21].

Part of the activities within this project, which then accept the standards promoted by UNESCO, is the construction of innovative solutions for establishing dialogue and cooperation between the media and specialised fact-checkers. This would include activities such as supporting the development of specialised units for checking information within the existing media newsrooms, especially through education and workshops for the use of new tools and technologies for checking online materials, and supporting self-regulation such as the establishment of a media council that would include fact-checkers, thus accepting the application of professional journalistic standards in their work.

In any case, the aim of the information verification process carried out by independent fact-checking organizations is to educate the public, improve the quality of political behaviour and improve the quality of journalism (Amazeen, 2020).

The number of organisations engaged in information verification is constantly growing, and in recent years the processes and practices of information verification, besides manual identification of

19 https://www.washingtonpost.com/news/fact-checker/

20 https://en.unesco.org/news/strengthening-relationship-between-independent-fact-checkers-and-media-outlets-promise-quality

21 https://en.unesco.org/trust-in-media-see

inaccurate content and their correction, have been increasingly expanding towards computational fact-checking (see Chapter 4).

The Center for Journalism Research at Duke University in the U.S.A. has been mapping and monitoring the development of specialised fact-checking projects globally for years[22]. When identifying relevant projects and organisations, they apply criteria similar to the principles of the International Fact-Checking Network (IFCN), and their database contains active and inactive projects, which are specifically recorded and counted. Also, they try to record the status of organisations that periodically check information during special events, such as elections. So far, they have mapped a total of 356 fact-checking organisations in the world. In their latest report published in June 2021, it is claimed that specialised information verification organisations now operate in at least 102 countries, that is, in more than half of the world's countries.[23] However, after years of steady and sometimes rapid growth of specialised fact-checkers, the trend seems to be slowing down in spite of the fact that misleading content and political lies are still a major challenge.

Over a hundred fact-checking organisations from around the world are now members of the International Fact-Checking Network (IFCN), which operates within the Poynter Institute. The IFCN's Code of Principles amounts to the standard on a global scale. The IFCN was founded in 2015 with the aim to bring together a growing community of fact-checkers around the world and advocate for the global fight against disinformation.[24] All member organisations of the Network are committed to respect the Code of Principles in the process of verification of information. The code of principles includes a set of rules, i.e., values that organisations commit to respect if they want to be members of the IFCN.

These principles include:

- → Commitment to Nonpartisanship and Fairness;
- → Commitment of Transparency of Sources;
- → Commitment to Transparency of Funding and Organisation;
- → Commitment to Standards and Transparency of Methodology;
- → Commitment to Open and Honest Corrections Policy.[25]

In order to be able to join the IFCN, that is, to be a signatory of the Code of Principles, one must meet six conditions:

1. The applicant should be a legally registered organisation, or a special team or unit within a legally registered organisation, and details of this can be easily found on the organisation's website.
2. The team, unit or organisation is established solely for the purpose of fact-checking.
3. The applicant has published an average of at least one factcheck per week over the course of six months prior to the application date, or 12 months if operating in a country that is home to five or more existing verified signatories at the date of application.

---

22 https://reporterslab.org/fact-checking/

23 https://reporterslab.org/fact-checking-census-shows-slower-growth/

24 https://www.poynter.org/ifcn/

25 https://ifcncodeofprinciples.poynter.org/know-more/the-commitments-of-the-code-of-principles

4. On average, at least 75% of the applicant's factchecks focus on claims related to matters that, in the IFCN's view, relate to or may have an impact on the welfare or well-being of individuals, the general public or society.
5. The applicant's editorial output is not, in the IFCN's view, controlled by the state, a political party or a politician.
6. If the organisation receives funds from local or foreign state or political sources, it should publish a statement on its site setting out to the satisfaction of the IFCN, how it ensures its funders do not influence the findings of its reports.[26]

The IFCN believes, as it is stated on its website, that "nonpartisan and transparent fact-checking can be a powerful instrument of accountability journalism".[27] Fact-checking organisations that are part of the IFCN play an important role in overseeing the content posted on Facebook.

In 2016, Facebook launched a programme to verify the information published on the platform by involving "third-party" or independent IFCN-certified fact-checkers in its content monitoring (so-called *Third Party Fact Checking Program*).[28]

"Since we do not believe that a private company like Facebook should be the arbiter of truth, we rely on independent third-party fact-checkers to identify, review and rate potential misinformation across Facebook, Instagram and WhatsApp."[29]

Although Facebook states that its fact-checking programme is extremely effective, and a number of scholarly papers show that exposing disinformation has positive effects (e.g. Chan et al., 2017; Faesen et al., 2021; Walter and Tukachinsky, 2020) believe that platforms have not yet provided concrete evidence to unequivocally prove the effectiveness of their fact-checking programmes.

There are two fact-checking organisations in Croatia: Faktograf and AFP. Both are signatories to the IFCN Code of Principles and are involved in Facebook's *Third Party Fact Checking Program*. Faktograf was launched in 2015 as a joint project of the Croatian Journalists' Association and GONG. In 2018, GONG became an independent publisher of Faktograf, and from 2021, Faktograf has operated as an independent organisation.[30] The fact-checking programme of the French international news agency Agence France-Presse (AFP), as part of its network for Central Europe, operates in Croatia. AFP is also a partner organisation in the Adria Digital Media Observatory, which is a regional hub of the European Digital Media Observatory (EDMO) that includes Croatia and Slovenia.

The first independent fact-checking project in Croatia was implemented by GONG in partnership with the Faculty of Political Science of the University of Zagreb back in 2007. The project analysed the merits and factuality of the messages that the main Croatian politicians and parties communicated to the public during election campaign through various channels, from the media to ads.[31]

---

26 https://www.ifcncodeofprinciples.poynter.org/application/public/20-minutes-fake-off/47A3F519-D9C4-6EB2-8E2C-5F0D29E5DAD3

27 https://ifcncodeofprinciples.poynter.org/know-more/the-commitments-of-the-code-of-principles

28 https://www.facebook.com/journalismproject/programs/third-party-fact-checking

29 https://www.facebook.com/journalismproject/programs/third-party-fact-checking/how-it-works

30 https://faktograf.hr/o-nama/

31 https://gong.hr/2007/11/12/gong-adwatch-pracenje-politickih-poruka-tijekom-ka/

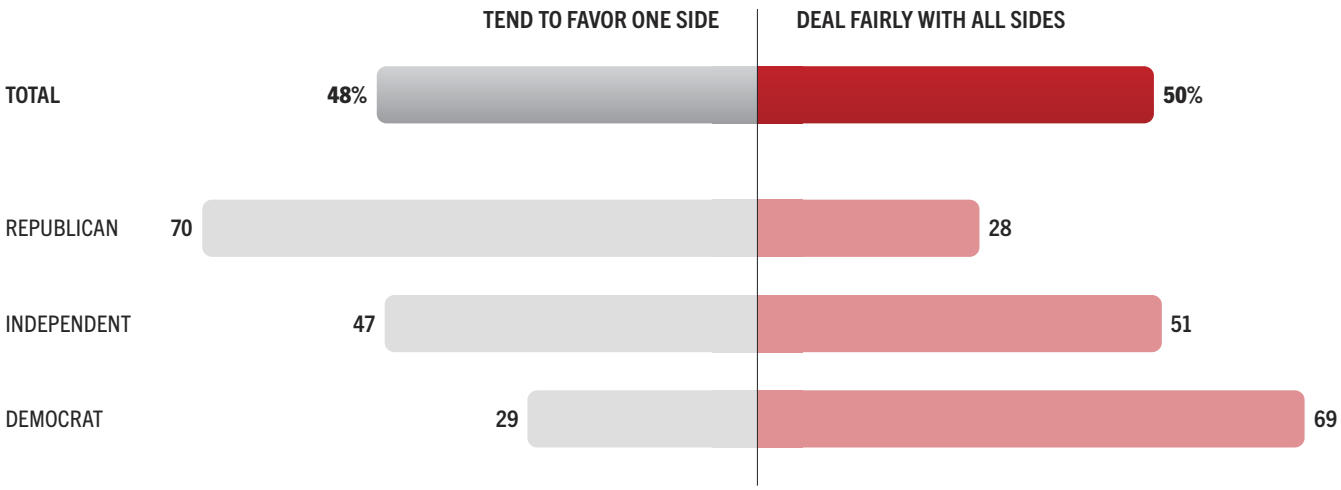
3.1. Problems of fact-checking and criticism

There is still relatively little research addressing the effects of activities carried out by fact-checking organisations, and the vast majority of research is still referred to the American context (Nieminen and Rapeli, 2019).

Fact-checkers and, generally speaking, the idea of establishing factuality encounter lots of criticism. The most frequent ones, as Louk Faesen et al. (2021) point out, are allegations of bias, low public support, and the question of who decides what can be checked. Faesen et al. (2021) argue that many accuse fact-checkers of favouring one worldview to the detriment of another. The 2019 survey conducted by the Pew Research Center in America found that the American public was divided in terms of the question of whether fact-checkers favoured one side or not – 48% of respondents claimed that fact-checkers favoured one side, while 50% felt they were treating both sides fairly (see Pew Research Center, 2019) (see Graph 2).

Americans split on whether fact-checkers favor one side; most Republicans say this is the case

% of U.S. adults who say fact-checking efforts by news outlets and other organizations:

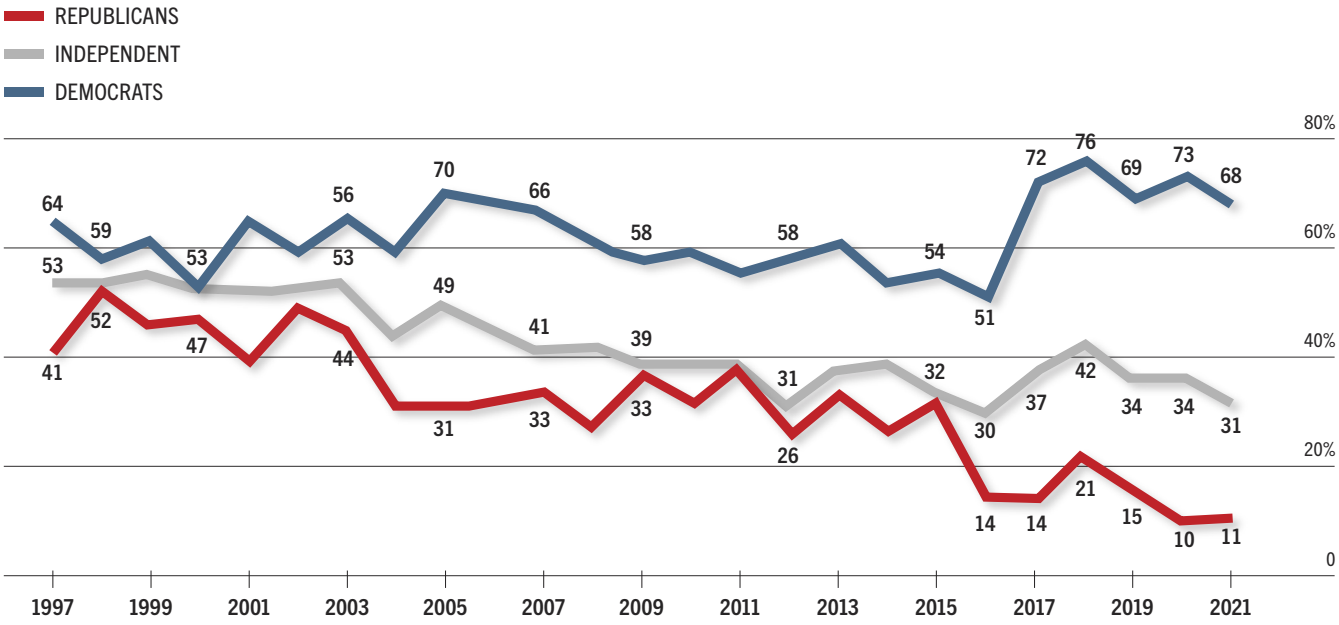


Graph 2: Perception of fact-checkers' bias. **Source:** Pew Research Center, February – March 2019

70% of the Republican Party supporters and voters think that fact-checkers are biased, while only 29% of the people who support Democratic Party think so. This gap overlaps with the level of trust in the media, which is extremely low among Republicans and was only 11% in 2021 (see Graph 3).

Americans' Trust in Mass Media, by Political Party

In general, how much trust and confidence do you have in the mass media – such as newspapers, television and radio – when it comes to reporting the news fully, accurately and fairly – a great deal, a fair amount, nor very much or none at all ?



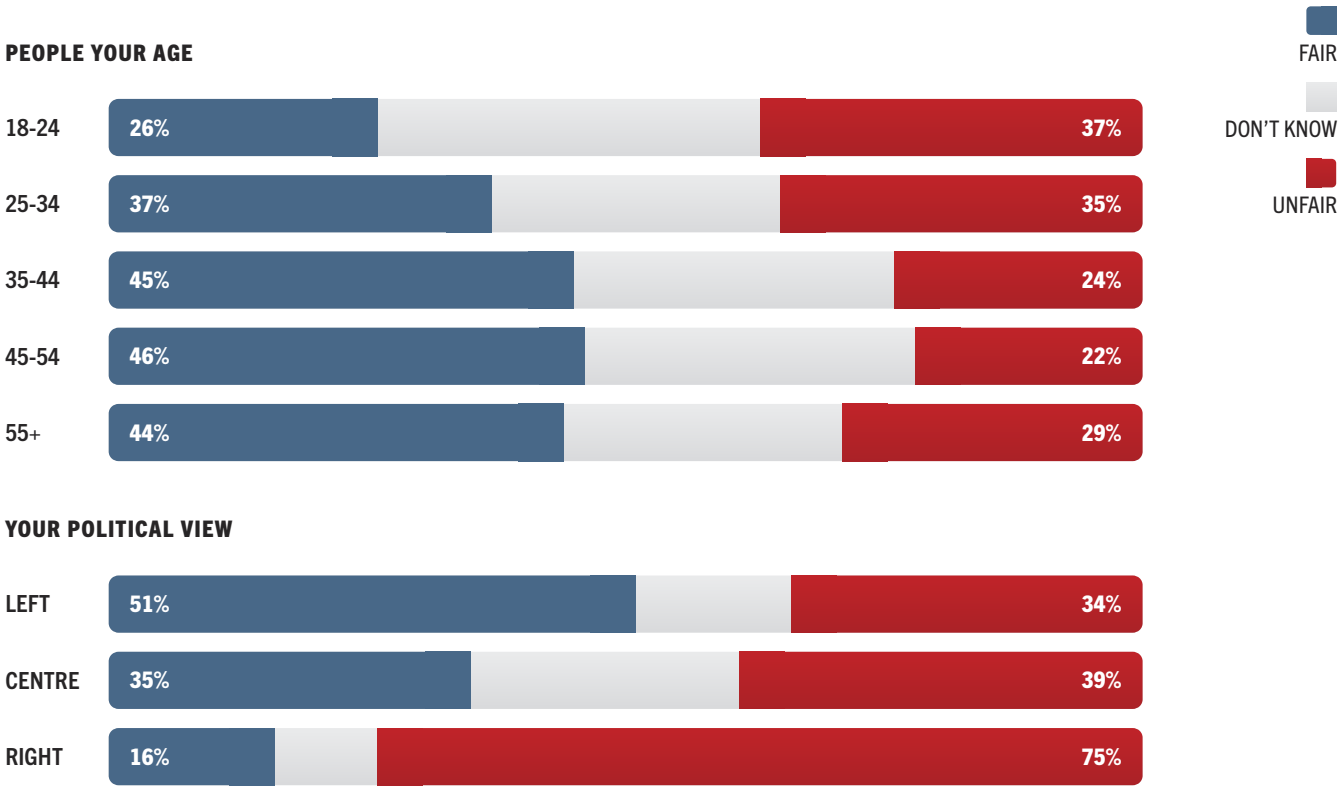
Graph 3: Americans' trust in the media. **Source:** Gallup, 2021

Faesen et al. (2021: 92) conclude that distrust of fact-checking organisations, which exists in one part of the population, is not great "for either factcheckers nor platforms. If platforms use factcheckers, a significant percentage of their userbase will be unhappy. Factcheckers themselves need legitimacy to work effectively: something that is becoming more and more difficult in increasingly polarized societies". Other researchers have reached similar conclusions. For example, Emily Saltz et al. (2020) analysed users' reaction to labels that Facebook assigns to COVID-19-related visual content that had been recognized as "misinformation". The authors found, among other things, that labels are perceived by many as "overly paternalistic, biased and punitive" (2020: 10). They believe that such system can "provoke feelings of distrust and hostility toward platforms and content correction labels". Faesen et al. (2021) similarly believe that many users perceive labels as censorship and restrictions on freedom of speech. According to Saltz et al. (2020), platforms, researchers and journalists should first of all seek to understand the feelings of discontent and uncertainty that push many towards misinformation in alternative media narratives. They also warn of the "backfire effect", which implies that those who believe a claim will believe it even more strongly if someone tries to convince them that what they believe is not true (2020). Although the authors did not identify this type of effect in their research, they still argue that "a defensive backfire reactions may be common in practice for American social media users encountering corrections on social media posts about news topics" (2020: 10).

On the other hand, Briony Swire-Thompson et al. (2020) argue that corrections published by fact-checkers are extremely unlikely to strengthen belief in misinformation. They cite the positive effects of fact-checking and urge fact-checkers to continue posting corrections. However, they warn that the presentation of corrected content should be tactful and should always include real, accurate information about what is marked as incorrect. A large number of empirical papers dealing with the effects of debunked or corrected information make similar recommendations for fact-checkers on how to approach debunking. Nathan Walter and Riva Tukachinsky (2020: 171) suggest that the correction will have the greatest effect if misinformation is responded to quickly, if the correction is compatible with the worldview of the recipient of the message or correction, and if the correction is attributed to the same source that was responsible for the misinformation. While the worldview requirement is quite difficult to achieve, the other two requirements represent a valuable recommendation for fact-checking organisations, as well as for the media whose role in correcting disinformation and misinformation is extremely important, especially if they themselves are the source of inaccurate information.

Building on the research that has established "ideological asymmetry" in the perception of fact-checkers in the U.S., where fact-checkers have less favourable image in the more conservative part of the population, which is associated with a low level of trust in the mainstream media among the right-wing electorate (see Graph 4), Ben Lyons et al. (2020) conducted a survey on attitudes towards fact-checkers in six European countries: Sweden, Germany, Italy, Spain, France and Poland.

Proportion that thinks the news media cover people their age and with their political views fairly
in United States of America



Graph 4: Americans' perception of balanced media coverage. **Source:** RIDNR, 2021

They found that the recognition and acceptance of fact-checkers in Germany and Sweden are much higher than in other researched countries. They also established that "anti-elitism" is an important predictor of negative sentiment towards fact-checkers. Likewise, the results of their research indicate that individuals who lean towards a left-wing worldview, who are sympathetic to the European Union and who are in principle satisfied with democracy, perceive fact-checking more positively. Lyons at al. (2020) conclude that political divisions are reflected in the perception of fact-checkers and that those who do not trust fact-checkers may be more receptive to disinformation that further deepens divisions.

Lately, and especially since the beginning of the Russian aggression against Ukraine, another problem has emerged – fake fact-checkers. This problem poses a threat to the very idea of verifying information. As the European Digital Media Observatory (EDMO) warns, the aim of such initiatives is to produce confusion and doubt about all the information that appears in the public space, so that it is increasingly difficult for citizens to assess what is true and what is false[32]. The EDMO cites the example of a Russian site that publishes some authentic corrections of incorrect information, but also publishes false corrections, i.e., "exposes" the footage allegedly launched by the Ukrainian side, while the footage was actually launched by the fact-checking site itself, which then allegedly 'corrects' this information. The goal of such perfidious manipulation is to confuse citizens so that they do not know who and what to trust, i.e., what is real and what is illusion.

32 More on **edmo.eu/2022/03/17/russian-propaganda-disguising-as-fact-checking-a-statement-from-the-edmo-taskforce/**

3.2. Fact-checking, public media, and news agencies in the fight against disinformation

The mission and purpose of the public media service has always been based on the values of universality, independence, public interest, pluralism, quality, and accuracy. Such a special status and role of public media services entails their special responsibility in the fight against disinformation, both through high-quality and responsible journalism, as well as through specialised programmes and projects for the development of verification of accuracy of information, media literacy and raising the general level of public awareness about this problem and available solutions. It is believed that the mission of the public service is to influence public discourse and increase the reach and influence of credible information. In addition, the effect and reach of what specialised fact-checking organisations do depend on their relationship with the established media (Graves and Cherubini, 2016).

In Europe, there are several examples of public media services that develop special fact-checking departments, apply new skills and new technological solutions, popularise verified and accurate information through their own media and online platforms, and participate in media literacy projects. In the UK, two public broadcasters – the BBC and Channel 4 – have set up their own fact-checking services, which have consolidated their status in recent years, as a response to the rise of information challenges. The BBC's RealityCheck started operating in 2015. In 2016, it strengthened its activities ahead of the Brexit referendum, while in 2017, it established a permanent editorial team (Graves and Cherubini, 2016; Samuels, 2017). In 2005, Channel 4 launched FactCheck in the form of a blog concerned with establishing the factuality of statements in general elections. This blog has been acting as a permanent department since 2010 (Graves and Cherubini, 2016). Finland's public service Yle, which has maintained its central role in the Finnish media system despite the proliferation of commercial and online media, pays special attention to exposing disinformation (Horowitz et al., 2021). Yle has produced a number of programmes and documentaries dealing with information disruption. In addition, Yle is very active in the field of media literacy development among different generations: from creating educational and interactive content for children and young people, to advise on digital life and information for senior citizens.

In Norway, the public media service NRK partnered with commercial media: TV2, VG and Dagbladet, in the creation of a common fact-checking platform – Faktisk. VG and Dagbladet were previously competitors in the market, and this is their first journalistic collaboration[33].

Along with public media services and leading news agencies, they are redefining their role in the modern information environment. Agence France Presse (AFP) has had the AFP Fact Check newsroom since 2017. It is specialised in fact-checking, with a global network of journalists covering online content in different languages, taking into account local cultures and politics, and working with AFP offices around the world to investigate and refute false, harmful and manipulative content. As mentioned before, AFP's fact-checking network includes Croatia as well. The German news agency DPA (Deutsche Presse-Agentur) also has a newsroom specialised in verifying information. It has recently launched the "Faktenscheck21" project to train journalists in fact-checking during the election year. The project is also implemented through an open digital learning platform and Slack channel for continuous exchange among members of the German fact-checker community. Both AFP and DPA are signatories of the IFCN's code of principles and both work with Facebook as independent verifiers of information circulating on this social network.

---

33 https://faktabaari.fi/avoinyhtk/why-faktisk-no/

4. Computational fact-checking systems

4.1. Defining key (technical) concepts related to technology and artificial intelligence

Computational approaches for fact-checking are largely based on Artificial Intelligence technologies. Artificial intelligence (AI) is a part of computer science that focuses on developing the ability of computers to perform tasks that require some form of intelligence, such as adapting to new situations, learning new concepts, deriving logical conclusions, understanding natural language, or analysing visual scenes. Machine Learning (ML) is a key component of AI technologies and a branch within AI. ML deals with developing algorithms that improve their performance based on empirical data. There are several categorisations of machine learning methods. One of the basic ones is the categorisation into supervised and unsupervised ML. Supervised machine learning uses datasets typically labelled by human annotators while unsupervised machine learning develops methods that are able to find patterns within unlabelled datasets.

Datasets are sets of examples used in machine learning models. For example, they can be comprised of text data containing comments from social networks or image data, which contain photos found in newspaper articles. Datasets can be labelled or unlabelled, e.g., a comment from social networks can be labelled as true or false. Datasets containing multiple types of data are the basis for the development of multi-modal models. Such models exploit the information contained in sources from diverse modalities, provided they contain complementary contextual clues. For example, a claim in the text may be interpreted as sarcasm on the basis of an attached photograph whose content is misaligned with the text.

Dataset labels are assigned by a labelling (annotation) process, as it is most often denoted in computer science terminology, or a coding process, which is how it is more commonly referred to in social sciences and humanities..

There are many specialised branches of research within the AI field, and one of the most important ones is Natural Language Processing (NLP), which combines machine learning methods with knowledge of linguistics and other related fields and is engaged in research into procedures for computational processing of data in natural language (in most cases, yet not necessarily, in the form of text).

Most of the data used in computational fact-checking systems are stored in text form. For example, in a text, you can find claims, i.e., factual propositions that can be true and false. One of the tasks within NLP, which uses datasets consisting of claims, is to detect where the claim is in the text (claim

detection). To this end, more general NLP methods can be used, such as determining semantic textual similarity (STS), which aims to establish the content similarity of two parts of a text.[34] Another important task of computational fact-checking is to establish claim check-worthiness which is aimed at identifying those claims that have broad social implications and thus a greater priority for verification. One of the NLP tasks that can help experts make a good judgment about a claim is to disclose and collect evidence, i.e., a set of data and sources that are useful in the process of verifying the veracity of the claims. For example, evidence can be text sources, knowledge bases, or sets of previously verified claims. The task of evidence verification is an NLP task aimed at verifying the credibility of evidence. All the NLP tasks listed so far actually constitute support for the NLP task of fact-checking, i.e., factual verification, or claim validation, which is concerned with determining whether a particular claim is true based on credible evidence, i.e., whether that claim / piece of information is actually a fact.

Models developed for individual computational fact-checking tasks often rely on models developed within other NLP tasks that have wider applications. One such task is named entity recognition (NER), i.e., recognising entities in a text, such as people, organisations, locations, dates and currencies. For example, the output of the model for NER may be to identify London as a city within a claim. The identified named entity can then be stored in a structured form of a so-called triplet in a knowledge base such as (London, Capital City of, England). Knowledge bases (KB) contain world knowledge in a structured form suitable for computer processing.

One of the applications of such structured data can be to enhance model interpretability, the focus of an NLP task that is aimed at explaining why models made a certain decision or prediction. Furthermore, the possibility of computer-aided enrichment of datasets with structured information can also have significant uses when applying the methods of Social Network Analysis (SNA), a related area which exploits the structure and interaction of various forms of social networks. SNA is an area of research that is located at the intersection of computer science and sociology, and uses algorithms and methods developed in many different disciplines. In the context of computational fact-checking, SNA methods have proved to be extremely valuable in identifying non-credible sources of information, both news portals and bots on social networks.

---

[34] For example, if we find that a part of a text is very similar to what we have previously known to be a claim, it is quite likely that this part of the text is also a claim.

4.2. AI technologies in information verification systems: literature overview, current experiences, and examples of good practice

4.2.1. A typical AI fact-checking system

The rapid development of computer-aided fact-checking systems was driven by the development of AI technology and the emergence of challenges associated with modern ways of communication such as the increasing influence of social networks on social processes. Over time, the key topics and tasks of interest within the field of computer-aided fact-checking have been identified, both by the academic and research communities, as well as by persons and organisations interested in implementing such systems.

An extensive overview of the use of AI technology for machine-assisted fact-checking can be found in several review papers dealing with this topic (Guo et al., 2022; Lazarski et al., 2021; Thorne and Vlachos, 2018; Zeng et al., 2021). Most of the current technology relies on natural language processing. Below is an overview of typical NLP tasks used in AI fact-checking systems, datasets and sources used, as well as an overview of the projects completed so far. A brief overview of general NLP techniques dealing with the processing of certain aspects of text (such as linguistic style), structured information (such as named entities), or unstructured information (such as frequent topics or clustering of similar texts) is given. In addition to fact-checking, this information can be used in additional research and analysis, for example in analysing the dissemination of claims through social networks using SNA-based methods.

The typical steps of a computational fact-checking system are described below. It is important to note that this is only a general description and that in different systems some steps may be omitted or interlinked differently. In practice, the steps described below are often implemented as machine learning models.

CLAIM DETECTION – the task is aimed at detecting verifiable claims in data (e.g., "VAT has risen by 1% since September"). This procedure is usually not fully automated, but finds probable candidates for such claims, which are then manually confirmed by experts. The input to this process is represented by the resources described below in section *Sources of claims to be verified*, and the output is a set of claims that need to be verified. An optional additional step is the task of discovering claims worthy of verification, which deals with the assessment of which of the claims found have a greater potential impact on society, and thus a higher priority for faster verification.

FACT-CHECKING / VERIFICATION OF CLAIMS – the task is to determine whether a given claim is a fact, i.e., whether it is true in view of the available set of evidence. Possible sources of evidence are described below in section *Sources of Evidence*. There are two main variants of this procedure. First, if the evidence is in the form of previously manually verified claims, an attempt is made to find an identical already verified claim for which the factuality is known. Second, in cases where an identical claim has not been found, elements (e.g., text documents) are identified from the set of evidence whose content might be relevant to determining the veracity of the given claim. Then, parts that are particularly relevant (such as sentences within documents) are extracted from these elements. Based on this information, a truth tag is automatically assigned. This designation can be binary (true/false) or on an ordinal scale (e.g., true on a scale of 1 to 5). Also, other categories such as "half-truth" or "There is not enough information to verify" can be added to the categories of truth/ falsehood. Finally, an important aspect of using ML models for these applications is their explainability. It is preferred to use explainable models that, in addition to deciding the veracity of the statement, also provide an explanation of which information most influenced that decision (e.g., which sentences, which words within sentences, which triplets from the knowledge base). Since such models are most often used semi-automatically, as an aid to a human expert, explainability makes it significantly faster and easier for an expert to determine whether the model has made a mistake.

4.2.2. Sources of claims to be verified

In most cases, statements can be automatically found in texts stored in digital form, but there are exceptions where this is difficult or impossible to achieve satisfactorily, for example from images or tables. The most common sources of claims are: (1) messages on social networks (Twitter, Reddit, Facebook), (2) Wikipedia, (3) online forums, (4) articles published on online portals (headline or body of the article, comments below the article), (5) debate transcripts, (6) online images (e.g., text within memes). Most of the sources are in English. However, there are also resources for other languages that are often limited in size. An example is the dataset described by Gupta and Srikumar (2021), which contains claims to be verified in as many as 25 languages.

4.2.3. Sources of evidence

The sources of evidence that can be used to computationally fact-check a given claim are primarily structured knowledge bases and unstructured text sources. A summary list of sources of evidence is provided below.

1. Structured sources - knowledge bases consisting of triplets, e.g., Google Knowledge Graph,[35] Other knowledge bases such as DBpedia (Auer et al., 2007), SemMedDB (Kilicoglu et al., 2012), topological properties of social networks (studied by SNA);
2. Unstructured text sources - previously verified statements from fact-checking pages (e.g., PolitiFact), scientific articles, Wikipedia, articles from credible online sources, specialised documents such as official public announcements or laws;
3. Metadata and other sources that may or may not be structured - tables, videos, images, publication dates, data about the source in which statements were published, or data about users who posted statements.

4.2.4. Related Areas/Tasks

There are numerous tasks in the field of natural language processing (NLP) that are conceptually related to the task of fact-checking or can be useful as an integral part of a computational fact-checking system. Below is a summary list of such tasks. **RUMOUR IDENTIFICATION** deals with the detection of rumours (unverified claims whose truthfulness is unknown) and the analysis of their spread through communication channels. **FAKE NEWS DETECTION** focuses on detecting published "news", which were deliberately created with the aim of providing the public with disinformation. **CLICKBAIT DETECTION** aims to find headlines of published articles that are not appropriately aligned with the content of the article with the aim of inducing users to access the article. **DECEPTIVE LANGUAGE DETECTION** is concerned with the discovery of texts that are written in a way that indicates dishonesty (e.g., deliberately vague statements). **AUTHOR PROFILING** deals with collecting the data on authors of texts with the aim of creating a profile of an author that can be used as additional information in establishing facts (e.g., an author who has published false statements in the past is more likely to repeat it). **ARGUMENTATION MINING** focuses on automatic recognition and extraction of argumentative structures from text data. **COMMON SENSE REASONING** develops computer models which are capable of generating new information by logical reasoning on the basis of previously known information. In the context of NLP, variants of this task (Poliak 2020;

[35] https://blog.google/products/search/introducing-knowledge-graph-things-not/

Storks et al., 2019.) are sometimes referred to as natural language inference (NLI) or recognising textual entailment (RTE). NLI/RTE can be useful in the context of a fact-checking task, if this task is formulated as verifying that whether the claim given is a logical consequence of the evidence given. **NAMED ENTITY RECOGNITION (NER)** deals with finding mentions of named entities in the text. **ENTITITY LINKING** deals with linking mentions of named entities in the text with structured information about them contained in knowledge bases. For example, "Angela Merkel" can be linked to the DBpedia base.[36] **COREFERENCE RESOLUTION** deals with linking references in the text that refer to the same entity (e.g., "Angela Merkel", "German Prime Minister", "She"). Most of these tasks are implemented in practice as supervised ML models.

4.2.5. Popular implementations

A relatively up-to-date list of tools used in the computational fact-checking process can be found at the following link: **https://www.rand.org/research/projects/truth-decay/fighting-disinformation.html**. Furthermore, leading IT companies have been recently adding tools to their offer that allow you to search for verified claims tagged with the ClaimReview scheme available on schema.org.[37] An example of this is Google's FactCheck Explorer (**https://toolbox.google.com/factcheck/explorer**)

1. **CLAIMBUSTER** (https://idir.uta.edu/claimbuster/)
   ClaimBuster is a publicly available online application that allows real-time information verification. It was developed at the University of Texas in collaboration with Meta (Facebook). Technical implementation is based on NLP and supervised machine learning methods, which means that models are trained on datasets labelled by humans.

2. **FULLFACT** (https://fullfact.org/about/automated/)
   FullFact has developed its own computational fact-checking system that works on data from news portals, social networks, but also speeches broadcast on public television channels. The system consists of a group of advanced tools that, for example, allow identifying claims worth checking or identifying people who are known to spread false claims. The tool is available to journalists, fact-checking organisations and emerging companies on request. FullFact is an example of an organisation that has achieved a high level of development of a computer tool for fact recognition by funding itself through open competitions, i.e., through *Google's Digital News Initiative*. Moreover, they follow the other two important guidelines for the development of a fact-checking system: (1) they communicate system limitations and flaws and (2) they use the ClaimReview scheme that allows easy integration of their data into other systems.

[36] https://dbpedia.org/page/Angela_Merkel

[37] https://schema.org/ClaimReview

4.2.6. Popular international NLP (Natural Language Processing) workshops on fact-checking

SemEval is a set of workshops in a competitive format, focused on solving individual NLP tasks related to semantic text analysis. Some of these tasks are related to computational fact-checking. For each task, the organisers of the competition provide a set of (usually labelled) data and infrastructure for evaluation. Groups of researchers participating in the competition develop a research prototype of a system that solves a given task in the short term and publish a publicly available system description paper about the developed system. Such workshops encourage a quick examination of different approaches to problem solving, which often results in useful knowledge that can be used in the development of production systems. The following is a list of relevant SemEval tasks through the years.

2021:
Task 6: Detection of Persuasion Techniques in Texts and Images (Dimitrov et al., 2021)
Task 9: Statement Verification and Evidence Finding with Tables (Wang et al., 2021)

2020:
Task 11: Detection of Propaganda Techniques in News Articles (Da San Martino et al., 2019)

2019:
Task 4: OtkHyperpartisan News Detection (Kiesel et al., 2019)
Task 7: Determining Rumour Veracity and Support for Rumours (second edition) (Gorell et al., 2019)
Task 8: Fact-checking in Community Question answering forums (Mihaylova et al., 2019)

2017:
Task 8: Determining Rumour Veracity and Support for Rumours (Derczynski et al., 2017)

Another popular workshop that has been held in parallel with scientific conferences in the field of NLP for 5 years in a row is "The Fact Extraction and VERification" (FEVER[38]). Similar to SemEval, it includes a competitive component in which researchers can solve fact-checking tasks. The description of the first edition of this workshop can be found in Thorne et al., 2018.

In addition to NLP conferences, workshops whose topic is computational fact-checking are also held as part of conferences in areas focused on more general phenomena in communication over the web and social networks. One such workshop is "De - Factify"[39] which is aimed at fact-checking in multimodal data (e.g., image + text), and is held in parallel with the "Association for the Advancement of Artificial Intelligence" (AAAI) conference.[40] Another such workshop is MEDIATE,[41] which is held as part of the "International AAAI Conference on Web and Social Media" (ICWSM).[42] Unlike the previous ones, this workshop does not include competitions, but only the publication of research papers on the topic of computational fact-checking, lectures and discussions.

38 https://fever.ai/
39 https://aiisc.ai/defactify/
40 https://www.aaai.org/
41 https://digitalmediasig.github.io/Mediate2022/
42 https://www.icwsm.org/2022/index.html/

4.3. Problems in the development of production systems for computational fact-checking

Many computational fact-checking systems did not see wide usage due to a number of problems. A prominent example is that of The Washington Post's TruthTeller, which promised real-time computer fact-checking on television channels and went out of business shortly after it was put into production.

There are two main and interrelated problems: (1) the immaturity of technology and (2) managing the expectations that interested parties have towards the capabilities of computer systems. In recent years, machine learning technology, especially natural language processing, has made it possible to use the ML model for many purposes with efficiency that sometimes comes close to human abilities. On the other hand, although the technology is suitable for solving certain steps in the process of computational fact-checking, due to the intrinsic properties of the models, AI/ML will not be able to completely replace people in this process for a long time, if ever, because many results and their interpretation depend on the cultural, social and political context.

The prevailing attitude of persons and organisations with experience in developing fact-checking systems is that future systems should specialise in individual subtasks that we know of or are expected to be performed by computer systems at a level commensurate to the human level.

Examples of such subtasks are: monitoring the source of evidence, identifying claims that are worth checking, verifying claims in relation to the database of stored claims, and automatically publishing verified claims in a structured format. In the vast majority of implemented solutions, the systems are not fully automatic, but automatic solutions have the role of supporting human fact-checking professionals by speeding up and facilitating their work.

4.4. Analysis of past activities and the role of AI in the field of fact-checking in Croatia

The most common sources of claims to be verified in Croatia are social networks (e.g., Twitter or public pages and groups on Facebook), online forums (e.g., forum.hr), and news portals (e.g., Index.hr or 24sata.hr). However, claims to be verified can also be found/disseminated through other media such as television, radio, public rallies and debates, etc.

There are still no papers in Croatia that directly deal with some of the main tasks in the process of computational fact-checking. However, related research and tasks have been tackled. Bogović et al. (2021) use topic models to analyse the most pronounced topics in the first nine months of the COVID-19 pandemic in news items and user comments on these news items. Similarly, Babić et al. (2021) study the ease of spreading news on Twitter, in relation to whether the news item is related to COVID-19 or not. Both papers use data in Croatian. Other studies conducted in Croatia on tasks related to computational fact-checking develop models for the English language because many more data are available for model training in English, which allows direct comparison with similar research in the rest of the world. Particularly relevant is the study by Boltužić and Šnajder (2016) dealing with the discovery of predefined claims in the text. In addition, the related tasks that were worked on was detecting events in the text (Di Buono et al., 2017b; Glavaš and Šnajder, 2013; Glavaš and Šnajder, 2014) and discovering hidden goals in the news using topic models (Korenčić et al., 2015). Likewise, research has been done with the aim to determine whether the claim was interesting enough to be news (Di

Buono and Šnajder, 2017) and to determine the intention behind the headlines of the article (e.g., surprise, quick spreading of news, shock, etc.) (Di Buono et al., 2017). Luttenberger et al. (2018) are concerned with detecting exaggerations in understanding claims from scientific literature. A popular task is also to detect propaganda in online news articles (Almer et al., 2020; Barišić et al., 2020; Pušelj and Škalec, 2020), then, similarly, revealing ideologically biased news they deal with (Palić et al., 2019). Finally, Anić et al., (2020) are focused on the detection of Twitter posts written by bots.

4.4.1. Language technologies for the Croatian language

In the context of fact-checking, most of the tasks come down to developing an ML/NLP model. These can be, for example, models for identifying named entities, detecting claims, measuring the similarity of claims with previously verified claims, etc. Common to all these models is that the text needs to be prepared before it can be used to train ML models. For this purpose, we need the language technologies described below.

4.4.2. Pre-processing

Some ML/NLP models require text pre-processing. For the Croatian language, the most important steps of pre-processing are morphological normalisation (Agić et al., 2013; Šnajder et al., 2008) and parsing (Agić and Merkler, 2013). Free tools are available for these tasks.[43] Programming libraries have been built around these tools that speed up and facilitate their use, e.g., Classla (Ljubešić and Dobrovoljc, 2019)[44] or Udpipe[45]

4.4.3. Vector representations of words

A popular approach to text analysis is to use vector representations of words in which similar words have similar representations (although they may sound completely different, e.g., plane and aircraft). There are many techniques for building such representations. Among the most famous ones are word2vec (Mikolov et al., 2013), GloVe (Pennington and Socher, 2014) and Fasttext (Bojanowski et al., 2017). The only inputs for these techniques are large amounts of text in the language for which word representations should be generated. For some of the techniques, there are already pre-generated vector representations for words in the Croatian language (Grave et al., 2018; Ulčar and Robnik-Šikonja, 2019). For the remaining techniques, tools[46] for building vector representations of words and large corpora of texts in Croatian, required for such tools to work, are available free of charge (Ljubešić and Erjavec, 2011). Vector representations of words are the basis for the development of specialised models of NLP.

4.4.4. Context-dependent vector representations of words

These are the models that are also trained on large volumes of text but represent an improvement over static vector representations of words. They allow generating a vector representation of a word that is not fixed but depends on the context in which the word is used. For example,, the

43 http://nlp.ffzg.hr/resources/models/

44 https://pypi.org/project/classla/

45 https://ufal.mff.cuni.cz/udpipe

46 https://nlp.stanford.edu/projects/glove/, https://code.google.com/archive/p/word2vec/

Croatian word "list" can be a fish (Eng. *sole*) or a sheet of paper, depending on the sentence. These models allow the development of specialised NLP models better than those that use fixed word representations and represent the current state-of-the-art technology in the field of NLP. Examples of such models are Elmo (Peters et al., 2018) or BERT (Devlin et al., 2018). There are several such models for the Croatian language. The BERTić model (Ljubešić and Lauc, 2021) was trained on texts from the web. A similar case is the cseBERT model (Ulčar and Robnik-Šikonja, 2020), which simultaneously includes Croatian, Slovenian and English. Finally, the Cro-CoV-cseBERT model (Babić et al., 2021) is an upgrade to cseBERT that includes additional training on texts related to the Covid-19 pandemic.[47] All of the above models are publicly available.

4.4.5. Projects and software solutions

ProFact is a project aimed at education and fact-checking in the context of disinformation about COVID-19 in Croatia through multidisciplinary research, raising public awareness and increasing verification capacity. More information can be found on the official website of the project: ***www.fpzg.unizg.hr/znanost_i_istrazivanja/projekti/medunarodni_projekti/pro-fact***

InfoCoV: Multilayer Framework for the Information Spreading Characterization in Social Media During the COVID-19 Crisis is a project of the Department of Informatics of the University of Rijeka that aims to study communication related to the coronavirus pandemic on social networks by means of multiple levels of analysis using modern language technologies for the Croatian language and social network analysis tools. The project is particularly interesting in the context of computational fact-checking due to the large number of disinformation that emerged on social media during the pandemic. More information can be found on the official website of the project: ***https://infocov.uniri.hr/***

Embeddia is a project that an international consortium is working on with the aim of developing NLP models that are applicable in several languages targeting primarily the languages of the European Union, including Croatian. Among the results of this project is the already mentioned language model cseBERT (Ulčar and Robnik-Šikonja, 2020) as well as a model for detecting offensive user comments in Croatian (Pelicon et al., 2021). More information can be found on the official website of the project: ***https://embeddia.eu/***

TakeLab Retriever is a project developed by the TakeLab[48] laboratory at the Faculty of Electrical Engineering and Computing in Zagreb. The aim of the project is to build a system that collects, indexes and processes newspaper tests available on Croatian news portals since their inception by means of advanced NLP methods. The system is still under development and is not open to the public, but it is already used for the purpose of several scientific projects. Although there is no implemented functionality of computational information verification, it uses some of the text processing methods used in fact-checking such as recognising named entities and linking them and has the ability to collect data from portals where incorrect information is likely to be found. TakeLab News Explorer is a project forerunner of TakeLab Retriever that did not see wide usage, but which used similar functionalities with technology that was available 15 years ago.

47 This model is particularly suitable for the development of specialised models intended to work on data related to the COVID-19 pandemic.

48 **https://takelab.fer.hr/**

4.5. Suggestions for possible activities and projects related to the use of fact-checking technologies and basic criteria, including the applicant's competence and capacity as well as the evaluation criteria

The common framework for the development of activities and criteria includes:

1. creating production systems that will perform their primary function of information verification;
2. developing professional competences for the development, refinement and maintenance of such a system;
3. enabling quality scientific research and submitting scientific project proposals;
4. developing basic AI/ML technologies for the Croatian language that can be used for other purposes.

Project evaluation should consider (1) the necessary competencies involved in the development of fact-checking systems, (2) defining criteria for scoring applicants to project-specific tenders, and (3) defining the impact level of the project on capacities to develop information verification systems. The required competences of specialists largely depend on the tasks covered by the project. For the development of language technologies, it is necessary to involve the scientific community that deals with areas of computer science related to AI technologies such as natural language processing. Participants are expected to have scientific titles or to be PhD students in the field.

For content analysis and cooperation on developing systems that include components for which only technical knowledge is not sufficient to solve, it is necessary to involve the wider scientific community.

For the technical implementation of software solutions that include databases or network applications, it is necessary to involve software engineers and other experts specialising in individual system components. For example, frontend and backend engineers or customer experience specialists.

The criteria for scoring in the evaluation should consider the capacities of the applicant team in terms of the previously mentioned competencies of experts necessary for the development of the system and past experience in the implementation of similar projects. Particular attention should be paid to the experience in developing AI solutions due to the additional challenges they bring.

The definition of the value of the proposed projects should also depend on the level of impact on society at different levels. Roughly, these levels can be divided into: (1) technical, (2) human, (3), scientific and (4) others.

Technical impact refers to the defined procedures and knowledge necessary to complete sustainable projects, the availability of solutions to the wider community that can build on them and propose new solutions, and the expected specificity of developed technological solutions, with emphasis to be laid on technologies that are applicable for other purposes.

Human impact refers to the expected level of competence development of all involved in the development of solutions and the level of knowledge exchange, with the emphasis on the development of interdisciplinarity. Greater value should be assigned to projects that offer greater knowledge transfer between different professions.

Scientific impact refers to the impact of scientific results at the international level, raising the competences of the involved scientists, establishing cooperation between different scientific institutions, both with each other and with external participants, and creating resources that can be used in further scientific research such as labelled datasets.

Other influences include a broader impact on society, such as developing awareness of the possibilities and benefits of applying AI tools for fact-checking purposes, but also for other similar tasks.

Even though specific aspects of evaluation and examples of impact primarily depend on the objectives of the projects, below are examples of some aspects that could be considered for the development of competencies in the context of the development of computational fact-checking systems. The most important of them is probably the cooperation of different professions on the development of individual modules. Experts of different professions can better analyse different aspects of complex problems that include rapid information change, analysis of large amounts of data coming from a variety of sources that need to be identified, and topical and conceptual differences between these sources that need to be contextualised.

The tasks of a computational fact-checking system imply a specific need for quality evaluation, such as identification of bias in data and models, requirements for explainability of model outputs and the possibility of realistic assessment of system performance in relation to professional fact-checkers. For this reason, it is necessary to choose suitable metrics of the efficacy of the system when defining projects.

Fact-checking systems are also specific for the large engagement of qualified fact-checkers in terms of labelling different datasets. When defining projects, it is necessary to pay special attention to the involvement of people in the key components of the system (human-in-the-loop) and to devise ways in which the labelling process can be facilitated or accelerated and additionally evaluate projects that include the development of solutions that enable this.

4.5.1. Activities

It is necessary to encourage cooperation among all professions and to build a knowledge base that enables the understanding of the problems of computational fact-checking in as many people as possible.

Examples of concrete activities at the technical level:

1. development of individual modules within the information verification system (for example, detection of claims, finding evidence, recognition of bots, profiling of authors, analysis of social interactions, explanation of models, etc.);
2. development of a web application that allows the use of developed modules through a unified and open programming interface (API).
3. development of data collection systems from various sources with support for open standards such as ClaimReview, computer translation into Croatian and the ability to connect to the collaborative platform;
4. development of language technologies for the Croatian language of production-level performance (for example, language models, models for resolving co-references, for identifying and linking named entities);
5. developing data labelling tools with advanced techniques to reduce the number of examples needed to train machine-learning models (using active learning methods for example) and with appropriate labelling quality metrics supported;
6. development of tools for continuous retraining and evaluation of information verification system models;
7. development of extensions for web browsers that allow the use of information verification systems.

4.6. Analysis of potential risks or possible weaknesses and problems in the use of technology and AI in the fact-checking system

The use of AI technologies in fact-checking systems has become irreplaceable due to many advantages like the ability to process large amounts of data coming at high speed from heterogeneous sources. Despite all the benefits, the use of AI technologies brings many risks. The risks can be roughly divided into (1) those that come from the challenges of the nature of AI systems and how they are implemented and used, and (2) those concerning the human factor and society's willingness to embrace AI technology including legal regulation.

The main risk groups are those related to the management of technical risks, human resources, public trust in the system and legal aspects.

Listed below are several such identified risks. However, it should be borne in mind that specific selection of AI models or specific tasks within the computational fact-checking system potentially lead to the expansion of the risk list. For this reason, it is essential to ensure a high level of expertise of the persons who propose, manage and implement projects.

4.6.1. Technical risks

The development of AI technology has experienced a sharp rise in the last decade, which entails many challenges that other branches of computer science and software development have encountered before and for which they have implemented processes to bypass them. The processes of development of AI technologies are increasingly catching up and finding solutions to some of the problems, and due to the speed of development, it is necessary to pay attention to recent methods at the time of planning the development of concrete technical solutions.

4.6.2. Long-term sustainability and maintenance of the developed system

Machine learning methods, and especially deep learning methods, place extremely high demands on computational resources and require hardware such as graphics processor units (GPU), or so-called tensor processing units (TPU), many central processing units (CPU) and large amounts of memory (RAM). One of the most important decisions is the choice between acquiring the necessary computing resources or renting services in the cloud, typically on one of the larger cloud platforms (Google Cloud, Amazon AWS, etc.). Both options bring some pros and cons that can be classified under risks. The main advantages of cloud computing services are that they scale with needs, that there is no large initial investment in resources and that there is no need for highly qualified experts to maintain servers, which should be thought of when decisions concerning the acquisition of local server computers are made. The main advantage of local server computers is that they can be a more cost-effective option in the long run.

Also, in this case, higher levels of control are possible. When planning the development of the system it is certainly advisable to weigh the pros and cons of both approaches, both during the development of the system and for the period in which the system will work in production. In doing so, it should be borne in mind that AI systems constantly need to be maintained, i.e., that individual models within the system need to be retrained as new data arrive.

For example, due to the constant change in the use of language, new concepts and entities emerging or existing terms changing their meaning. This fact entails the need to secure computer resources during the entire period of use of the system. As models are trained on data coming from heterogeneous sources, changes in programming interfaces (API) and data format are expected. For this reason, it is necessary to provide procedures and processes that ensure the quality and availability of data by monitoring and aligning with changes in external systems. This entails ensuring technical and human capacities and constantly monitoring data quality and maintaining software for data processing and storage. Given the expected fluctuation of experts working on such systems, it is necessary to document processes, procedures, software, data and models. This is also necessary to detect the sources of errors in the operation of the system and enable them to be corrected. The development of AI models and data collection is a resource-intensive task, and it is necessary to establish good backup processes of all parts of the system so that the data is available to interested parties, while at the same time not having an impact on the operation and availability of the system.

When developing the system, attention should be paid to the security of potentially sensitive data and the access rights of participants in the development of the fact-checking system. It is necessary to provide clear procedures for ensuring control of access to various parts of the system, software repositories and datasets, as well as to make sure that best practices in protecting systems open to the public, such as online data entry applications or open programming interfaces (APIs), are monitored.

It is necessary to ensure monitoring of the system, logging of critical operations and access to certain parts of the system. It is necessary to establish procedures for ensuring the quality of data entry, i.e., ensuring the quality of the collected text, the quality of the labelled data, information on data labellers and the labels themselves. Such control allows the use of data for scientific purposes, allows the involvement of independent developers and provides additional basis for trust in the operation of machine learning models.

Also, the costs of collecting data should be borne in mind and included in the project's list of expenses. Some of the programming interfaces (APIs) are available free of charge to academic institutions, but often come with limited usage rights or technical limitations, such as the speed of access to data or limitations on the amount of data. This also applies to the API-s of some social networks as well as to repositories of certain newspaper sources.

4.6.3. Retaining and developing the expertise involved in system development

The development of a computational fact-checking system requires the specific expertise of the experts involved and it is necessary to make provisions and plan how to ensure that there is a sufficient number of qualified persons to design, develop and maintain such a system. The availability of experts is influenced by numerous factors such as a large fluctuation of employees in the real sector and academia due to high competitiveness, the fact that projects are limited by duration and funding, scarcity of experts with a certain academic degree, as well as specific knowledge necessary for the development of such projects. It is necessary to plan the development of the system in such a way as to include the various necessary technical competences for the development of individual parts of the system.

Although the core of a computational fact-checking system is developed by people trained to develop machine learning models, the final system consists of many other parts such as data storage systems, network applications consisting of different parts, such as background services and interfaces that must be installed and maintained. This entails planning to involve people with expertise in developing individual parts of the system, such as frontend and backend software engineers, engineers and designers for user interface development, database maintenance engineers, and DevOps engineers for infrastructure maintenance. Some experts may have several such roles, but this often entails a reduced quality of the final system.

4.6.4. Level of confidence in computational fact-checking system results

It is necessary to ensure the management of the expectations of the public, i.e., the persons who should use the developed system. To achieve this, it is necessary to transparently highlight the capabilities of the system at each level so that they are understandable to the expected audience. For example, one way to achieve this is to compare the performance of the developed system module with the performance of people who do the same job. When developing machine learning models, it is necessary to engage people who will monitor and evaluate the quality of system predictions and ensure that models are improved if they are found to differ from what is expected.

4.6.5. Legal aspects

RazThe development of AI technologies is extremely rapid and has a wide impact on society, which often leads to systems developing in an environment of uncertainty related to legislation and legal constraints.

It is necessary to ensure that adequate legal support is provided from the beginning of the development of the system to mitigate some of the risks associated with the use of AI technologies, in particular those related to data collection and processing. There are three main aspects to pay attention to:

(1) The GDPR imposes many restrictions, such as: which data can be used, how, for what purpose, and where can it be stored. It is necessary to ensure compliance with all regulations related to the GDPR,

(2) the copyright of journalists and private persons whose texts are used in the analysis,

(3) the right to access and use information obtained through APIs of various data collection services such as those from social networks (Facebook, Twitter, Reddit and others).

The last aspect is particularly challenging due to the frequency of changes in the rules of use, the increasingly strong regulations imposed on information providers and the great potential impact on the operation of the computer fact-checking system, to the point at which the collected data can no longer be used in the operation of the system

.

5.
Analysis of key documents, international and European standards, criteria and processes for establishing a fact-checking system

Disinformation is not a problem that can be tackled easily as it cannot be placed neatly within the existing branches of European and national policies. The problem itself is complex as it involves different actors, technologies, platforms, behaviours, motivations and content. Furthermore, it comprises the rapid pace of development of technology and online environments; the cross-border dimension of disinformation and operation of online platforms; the distribution of competences between the European Union and Member States in the creation of certain policies; national specificities, including legal, historical and socio-political ones, which affect the extent and form of the country's vulnerability to disinformation; and finally, the difficulty of clearly defining the boundary between disinformation and freedom of expression. European Union policies that try to address the problem of disinformation start from the premise that disinformation is not necessarily illegal, but is harmful (European Commission, 2020a; 2020b).

In 2018, the European Commission presented the foundations of its approach to combating disinformation. This approach was designed as a set of complementary actions in several areas; from strengthening media literacy and supporting quality journalism to increasing transparency and accountability of online platforms and protecting the privacy and personal data of citizens online.

The European Commission's action in this area was primarily driven by the scale of new technologies of political propaganda and manipulation witnessed ahead of the UK referendum on EU membership in 2016 and the US presidential election. By creating a comprehensive "European approach" to tackling disinformation, the European Commission primarily wanted to protect the European Parliament election in 2019 and democratic institutions in the European Union and its Member States. However, the approach was soon unexpectedly put to the test in information disorder related to the COVID-19 pandemic, and then in the context of Russia's war in Ukraine in 2022. This chapter presents the European Commission's anti-disinformation policy and the instruments and measures it encompasses. It also outlines the evolution of anti-disinformation policy from soft measures and self-regulation of online platforms towards co-regulation and horizontal legal solutions which try to ensure the transparency and accountability of tech corporations in the process of mediating public information.

The operations of online platforms are based mainly on their internal policies, business interests and technologies. In recent years, under pressure from the interested public and due to better resources and methodologies of the scientific community, the content moderation policies of leading platforms have become somewhat clearer, but still not transparent enough. For example, it is still unknown what criteria the algorithms of different platforms take into account when selecting, ranking and moderating different types of content to each individual user.

Barrett and Kreiss (2019) call the content moderation policies of leading platforms "transient." When they talk about "platform transience", they indicate that the corporate policies of platforms change over short periods, often dramatically and largely as a result of external pressures and economic reasons. These changes are so rapid that they are almost impossible to track, especially when one considers that leading platforms often apply different content policies in different markets in which they operate. For example, Facebook in the US presents their users news separately from other content, while in most European Union countries this is not the case and news is mixed with other content in the general feed. In addition, certain policies, although implemented in the same way, may have completely different implications in different cultural and political settings (Nenadić, 2020b).

Taking into account the complexity of the disinformation phenomenon, the European Commission first sought advice on how to deal with it by setting up a High Level Expert Group on Fake News and Online Disinformation (HLEG) in 2018. In March 2018, the Expert Group, consisting of 39 expert representatives of civil society, online platforms, media organisations, journalists and academia, produced a report proposing a multidimensional approach to tackling disinformation, based on five pillars designed to: (1) enhance transparency of online news, involving an adequate and privacy-compliant sharing of data about the systems that enable their circulation online; (2) promote media and information literacy, (3) develop tools for empowering users and journalists to tackle disinformation, (4) safeguard the diversity and sustainability of the European news media ecosystem, and (5) promote continued research on the impact of disinformation in Europe, which is critical for shaping evidence-based responses and policies.

This multidimensional approach became the basis of the European Commission's policy against disinformation, formally presented in April 2018 when the Commission published its *Communication to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions - Tackling Online Disinformation: a European Approach (COM/2018/236 FINAL)*, further operationalised through the *Action Plan against disinformation (J O I N (2 0 1 8) 36 final).* In both documents, the Commission refers to citizens' exposure to a large amount of disinformation online, including misleading and verifiably false information, as one of the main challenges Europe is facing. Even though media policies and the protection of the electoral process lie primarily within the competence of the Member States, and some Member States have independently set out to develop measures to protect electoral processes against online disinformation, the European Commission (2018) clearly states that "the cross-border dimension of online disinformation makes a European approach necessary in order to ensure effective and coordinated action to protect the EU, its citizens, its policies and its Institutions."

The European approach to combating disinformation online takes into account the complexity of the matter and the fast pace of developments in the digital environment so that it highlights the importance of a comprehensive and complementary response across several areas. This includes fostering media pluralism through supporting "high-quality journalism" and establishing a balance between those who create news (the media) and those who distribute it (online platforms); greater transparency and accountability of online platforms; investing in media literacy; encouraging the

development of an independent European network of fact-checkers and cross-border cooperation between different stakeholders in the process of detecting, analysing and combating disinformation. In addition, such approach emphasises the importance of harnessing new technologies and artificial intelligence (subject to appropriate human oversight) in verifying, identifying and tagging disinformation, which entails investing in systematic education of all those who will apply such technology.

In its *Communication*, the European Commission emphasises the importance of supporting "quality journalism" as "an essential element of a democratic society". It adds that "by ensuring a pluralistic and diverse media environment, they can uncover, counterbalance, and dilute disinformation" and encourages Member States to consider deploying tools to support the sustainability of "quality journalism" and innovation in journalism. *The Action Plan against Disinformation* contains a specific measure whereby Member States, in cooperation with the Commission, should scale up multidisciplinary teams of independent fact-checkers and academic researchers with specific knowledge about local information environments, to identify and expose disinformation threats on various social networks and in digital media.

With the arrival of the COVID-19 pandemic and the related 'infodemic' (WHO, 2020), the European Commission called for enhanced, coordinated and more transparent engagement of platforms through self-regulation. With a new Communication of 2020, entitled *Tackling COVID-19 Disinformation — Getting the Facts Right*, the European Commission took over the term coined by the World Health Organization, stressing that the infodemic associated with the COVID-19 pandemic is a global challenge and that collaboration with platforms is an essential element of an effective response. The Commission asked platforms to "deepen their work to combat the risks sparked by the crisis" and to report regularly on the measures implemented. One of the key actions agreed by the platforms at the initiative of the Commission was to promote and ensure greater visibility of information sourced by national and international health organisations and professional media. This is also in line with the requirements of the World Health Organization whose representatives met with representatives of thirty Silicon Valley companies in February 2020 to agree on a strategy to combat the infodemic. These pressures by international and intergovernmental organisations resulted in significant changes in content policies of the leading platforms. For the first time, platforms, as private companies, began evaluating content and sources based on "credibility". This shaped the information environment in the pandemic and may have far-reaching implications for the times after the pandemic.

5.1. EU Code of Practice on Disinformation

One of the key instruments of the European approach to tackling disinformation online is the Code of Practice on Disinformation announced in the Communication of April 2018 and presented in the autumn of the same year in order to be fully viable for the 2019 European Parliament elections. The Code was initially set up as a self-regulation by the leading online platforms, advertisers and advertising industry that have committed to: (1) improve the scrutiny of advertisement placements to reduce revenues of the purveyors of disinformation; (2) ensure transparency about political and issue-based advertising, by identifying sponsors and amounts spent; (3) mark automated accounts (bots); (4) empower users through the promotion of media literacy and providing greater visibility of trustworthy content; (5) enable the academic research community to access platform data so that it can track disinformation online and understand its impact. Some platforms — signatories to the Code — have formalised cooperation with fact-checking organisations that are now some of the key actors in the detection and labelling of content as false. The 2018 Code was signed by Facebook (including Instagram), Google (including YouTube), Twitter, Mozilla, Microsoft and Tik Tok. At the presentation of the Code, the European Commission announced that if such a self-regulatory tool did not prove effective, it would explore the possibilities of regulation.

The first years of application of the Code have shown several shortcomings of the approach which entrusts private actors with the important task of increasing transparency and credibility of the information environment. Above all, the platforms have not been sufficiently transparent to the research community and regulators, failing to provide functional access to data that would allow independent oversight of the implementation and impact of their activities.

In an Assessment[49] carried out by the European Commission after the first year of application of the Code, the Code was generally recognised as a valuable instrument in combating disinformation but containing some significant shortcomings. These include lack of uniform definitions among platforms (e.g., what exactly is implied by political advertising or disinformation on different platforms). Similar problems were highlighted by the European Regulators Group for Audiovisual Media Services (ERGA)[50] (whose member is also the Croatian Agency for Electronic Media). More specifically, the lack of uniform approach, the impossibility of independent and data informed monitoring and the impossibility of obtaining country-specific information. When explicitly requested by ERGA, the platforms did not provide functional access to the data necessary to understand their action against disinformation. Based on such experience, which makes it impossible to monitor the effectiveness of the Code, ERGA proposed to the European Commission to move from self-regulation of platforms to co-regulation that would include clearer principles of action and ensure public oversight of the effectiveness of the measures and activities of leading online platforms. The impossibility of independent oversight was a particularly significant shortcoming of the 2018 Code, as action against disinformation that is not sensitive to context, specific language or satire may result in restriction of freedom of expression, which is by no means acceptable in a democratic system in which freedom of expression is one of the fundamental rights. Therefore, deciding what freedom of expression is and what it is not, should not be left to private companies without public scrutiny. In addition, the 2018 Code brought together only a small number of platforms which did not include, for example, messaging applications such as Messenger or WhatsApp, although they have an increasing role in spreading disinformation (Elias and Catalan-Matamoros, 2020).

---

49 https://ec.europa.eu/transparency/documents-register/detail?ref=SWD(2020)180&lang=en

50 https://erga-online.eu/wp-content/uploads/2020/05/ERGA-2019-report-published-2020-LQ.pdf

Due to all of the above, in 2021, the European Commission initiated a process of amending the Code of Practice on Disinformation and published its Guidance on strengthening the Code of Practice on Disinformation (2021), notably through enhanced content and commitments and by introducing measurable targets and performance indicators and involvement of a wider stakeholder group among future signatories to the Code. The new, amended Code of Practice on Disinformation was signed and presented on 16 June 2022 by 34 signatories who participated in the amendment of the 2018 Code. The new signatories, along with leading online platforms and representatives of the advertising industry, include civil society organisations and fact-checking organisations, including the Croatian *Faktograf*.

Signatories have maintained key areas of action, such as: demonetisation of disinformation, ensuring transparency of political advertising, empowering users, strengthening cooperation with fact-checking organisations and ensuring access to data of leading online platforms for researchers. The new Code also contains a reinforced framework for monitoring the effectiveness of the implementation of agreed measures and announces the setting up of a Transparency Centre, which should provide the public with a clear and regularly updated overview of the policies and tools used by the platforms-signatories to the Code.

In addition, the 2022 Code focuses not only on disinformation but also on several measures against misinformation, which may result in platforms acting more frequently on media and journalistic mistakes (Nenadić, 2021). Even if "it is not an aim of the strengthened Code to evaluate the veracity of editorial content" (European Commission, 2021), this may still be one of its outcomes, especially considering how broad the concept of "editorial content" is and how difficult it is to define the "media" in the digital age. The 2022 Code, according to announcements, provides for greater engagement of fact-checking and other organisations in defining and applying tools for assessing the trustworthiness of sources (including media sources) which would then be given greater visibility in the algorithmic presentation of content on platforms signatories to the Code. Greater visibility also means a better negotiating position with advertisers who remain key sources of funding for media in the digital environment. Because of all this, the Code of Practice on Disinformation is one of the key instruments for the protection of informed citizenship, which also indirectly defines the status and sustainability of media in the online environment. Since social networks are often used by journalists as sources of information, but also as platforms for distributing or promoting their work, and given that every action of platforms in this area also affects the media, it is surprising that media organisations or associations of journalists are not part of the Code (as its signatories or in some other active form).

5.2. Digital Service Act (DSA)

The Digital Services Act (DSA), a landmark regulation for the protection of rights in the digital environment, entered into force on 16 November 2022 and will be directly applicable across the EU from mid February 2024. As regards the obligations for very large online platforms and very large online search engines, the DSA starts applying even earlier. The Act contains a set of rules requiring tech companies to properly assess and mitigate the harm their products may cause, as well as to make such assessments and harm mitigation measures available for scrutiny by independent auditors and researchers. In addition, the DSA includes a series of transparency commitments that are tailored to the type and nature of the digital service. Furthermore, the Digital Services Act updates and simplifies the existing system of flagging and treating illegal content on online platforms and contains specific measures to eliminate and combat the spread of illegal content from online platforms. Illegal content is, for example, hate speech or incitement to terrorism. Disinformation, while harmful, is not necessarily, and generally is not, illegal.

Disinformation is a very complex phenomenon that is intertwined with a fundamental right to freedom of expression. Therefore, the Digital Services Act does not apply removal measures to disinformation. In point of fact, it is not the same whether someone spreads false and misleading information intentionally to cause harm or make a gain, or unintentionally, naively believing that it might be true. Even the media can contribute to the spread of disinformation through lack of verification, sensationalism, clickbait headlines, etc. From the perspective of freedom of expression and freedom of the media, it would be problematic for social networks to have a legal basis for removing bad journalism. This is why the Digital Services Act promotes transparency, cooperation between different stakeholders and co-regulation as an approach to combat disinformation. More specifically, it envisages the transition of the Code of Practice on Disinformation, as an instrument to address the risk of disinformation, from self-regulation to co-regulation. That is why it is very important how the Code of Practice on Disinformation is structured, which signatories it covers and what measures it commits to, since the application of the DSA could strengthen the Code.

The Digital Services Act does not treat all platforms equally; instead, it requires the greatest level of transparency and accountability from the largest online platforms. Very large online platforms, according to the Act, are considered to be those that have more than 45 million users on the territory of the European Union.

5.3. European Democracy Action Plan (EDAP)

Almost in parallel with the presentation of the proposal of the Digital Services Act, the European Commission presented the European Democracy Action Plan (EDAP) in December 2020[51]. The purpose of this Action Plan is to build more resilient democratic societies in the EU by: 1) promoting free and fair elections, 2) strengthening media freedom, 3) countering disinformation.

In order to protect free and fair elections, the European Commission announced the drafting of a new law to ensure greater transparency in the area of sponsored political content (political advertising) on online platforms, which is often part of disinformation campaigns, as well as drafting accompanying measures and guidelines for political parties and Member States. In addition, it also announced a proposal to amend the Regulation on the statute and funding of European political parties and the establishment of a new joint operational mechanism within the European Cooperation Network on Elections to enable the deployment of joint teams of experts and improve work with the Network and Information Systems Security Cooperation Group in order to combat threats to electoral processes.

In order to strengthen media freedom and pluralism and provide journalists with a safer environment to do their job without pressure and intimidation, the Commission presented an initiative to combat the abuse of court proceedings (strategic lawsuit against public participation - SLAPP) and announced sustainable funding of projects for legal and practical assistance to journalists in the EU and beyond.

On 16 September 2021 the Commission presented its first Recommendation on the protection, safety and empowerment of journalists and other media professionals in the European Union.[52] One of the Commission's measures to support media pluralism is the new Media Ownership Monitor along with measures which consider models of support for diversity and sustainability of media, in particular the media of general interest. Due to the economic crisis caused to the media by the COVID-19 pandemic, the European Commission also warned of the importance of establishing measures for a transparent and fair allocation of funds for state advertising. This implies funds for media advertising of various initiatives and campaigns spent by various state bodies, ministries, institutions and enterprises in which the state has majority stakes. Such advertising can cover a wide variety of messages and topics in terms of its content, but since government and/or public funds are often a crucial source of financing for many media, it is critical that they are distributed on the basis of clear and fair criteria.

In combating disinformation, the European Democracy Action Plan provides for, inter alia, the establishment and strengthening of existing cooperation, research and action structures between independent regulators, the media, civil society, private sector actors and other relevant stakeholders. It also envisages an increase in support and financing of initiatives and new innovative projects to combat disinformation under various EU programmes, in particular those implemented by civil society organisations and higher education institutions, with the participation of journalists. The Commission also announced additional financing for the involvement of journalists in media literacy activities. The revised Audiovisual Media Services Directive requires Member States to promote the development of media literacy skills. In addition, it commits video-sharing platforms to establish effective media literacy tools and enhance user awareness. The development of media literacy is also promoted by the Media and Audiovisual Action Plan.

---

51 https://eur-lex.europa.eu/legal-content/HR/TXT/HTML/?uri=CELEX:52020DC0790&from=EN

52 https://ec.europa.eu/commission/presscorner/detail/en/ip_21_4632

5.4. European media in the digital decade Action Plan to Support Recovery and Transformation

In parallel with the presentation of the European Democracy Action Plan, the Commission also presented a Media and Audiovisual Action Plan[53]. The aim of the Action Plan is to accelerate the recovery, transformation and empowerment of the media industry in the EU, especially in terms of adjustment to the online market where most advertising revenue goes to global online platforms. Moreover, due to the pandemic and the context of economic uncertainty, advertisers have reduced overall advertising expenditure, further hampering the sustainability of the media sector.

Some of the approaches envisaged by the European Commission through this Action Plan include the following: easier access to EU support for all media companies, regardless of their field of activity and size; fostering cooperation and support for the information media sector; investing in the development of innovative media technology and education and cooperation related to its application. Start-up and scale-up media organisations will be able to participate in Creative Innovation Lab, a new initiative under the cross-sectoral component of the Creative Europe programme. It is designed to bring together the media sector and other creative sectors (e.g., music, publishing) and experiment with data, virtual and augmented reality, and other technologies to develop new content, business models and journalistic skills, as well as to promote inclusiveness and sustainability and encourage audience engagement. The Media and Audiovisual Action Plan also puts a great emphasis on the development of media literacy of citizens through programmes involving journalists and media.

53 https://eur-lex.europa.eu/legal-content/HR/TXT/HTML/?uri=CELEX:52020DC0784&from=EN

5.5. European Digital Media Observatory

At the core of the **European approach** to tackling disinformation is cooperation between different actors at national and European level, as well as multidisciplinarity of responses. This is why the European Digital Media Observatory (**EDMO**) was established in June 2020[54]. Housed at the European University Institute of Florence, EDMO is based on the work of an independent multidisciplinary community of researchers, fact-checkers and media literacy experts who, in collaboration with media organisations and online platforms, try to better understand the problem of disinformation and strengthen the resilience of society to this problem. The Observatory has filled a gap that existed prior to its foundation in the **European approach** to tackling disinformation. This approach calls for the fight against false information and manipulation to be multidisciplinary and to involve different stakeholders. EDMO is therefore designed as a platform that brings together relevant organisations and individuals[55] through five main activities so that they can pool resources and existing knowledge, share experiences, tools and instruments in fact-checking and developing media literacy, and create a secure harbour for researchers to access data from online platforms. This should contribute to a better understanding and analysis of activities, trends and techniques of spreading lies and manipulation online.

EDMO activities are as follows:

- → to run a secure online platform supporting the academic analysis of disinformation campaigns and trends on the one hand, and to provide public information to raise awareness about disinformation and manipulative online information on the other;
- → to establish a framework for **access to online platform data** for research purposes (platform data is necessary to understand the problem of disinformation spread and as such should be available to regulators and the research community in compliance with the General Data Protection Regulation (GDPR));
- → to support the coordination of independent fact-checking activities (fact-checking projects), including mapping out of fact-checking projects in Europe, and creating a directory aggregating fact-checks, and media literacy material from external repositories;
- → to support the coordination of academic research on the phenomenon of disinformation in Europe and to create a repository with relevant peer-reviewed scientific literature;
- → to conduct anti-disinformation policy research and analysis.

EDMO is managed by the **Executive Board** composed of representatives of partner institutions[56], who are also coordinators of various activities. The work of the Executive Board is supervised by the **Advisory Board**, composed of experts who provide strategic guidance to the initiative in line with the latest trends. The project is funded by the European Commission and its main partners are the European University Institute in Florence, the Danish Aarhus University, ATC - Athens Technology Centre and Pagella Politica, Italian fact-checking project. EDMO has both a European and a national dimension. The central EDMO and the multidisciplinary community gathered around it provide support to national research and media literacy campaigns, with the aim of strengthening the ability of citizens to assess the quality and veracity of information on the Internet. During 2021, the first national and

54 **https://edmo.eu**

55 **https://edmo.eu/meet-our-team/**

56 European University Institute, University of Aarhus, Athens Center for Technology, Pagella Politica **https://edmo.eu**

multinational EDMO hubs[57] were established, aiming to detect and analyse disinformation campaigns at the national level, analyse the impact of these campaigns on society and democracy, promote media literacy activities and monitor the policies of online platforms and the digital media ecosystem, in cooperation with national institutions. As of late 2022, the fourteen national or multinational hubs cover all 27 EU Member States as well as Norway, in the EEA. Local EDMO hubs workwith the central EDMO, but also with each other, to share best practices, relevant insights and useful content.

EDMO regularly provides both residential or online training modules on understanding and countering disinformation online. These multidisciplinary and interdisciplinary workshops are often tailored for journalists, fact-checkers, civil society organisations, policymakers and anyone participating in media and information literacy development programmes.

EDMO has also established a digital collaboration platform that brings together fact-checking organisations from different European countries that must meet certain criteria[58] and working principles. The EDMO platform enables them to communicate and share information, tools, services and resources to facilitate the detection and analysis of disinformation, especially the disinformation with a cross-border dimension. Given that computational fact-checking may never be entirely possible in all cases and that human judgment will always be necessary, the EDMO platform aims to assist its users in various fact-checking tasks that they perform, by optimising workflow, facilitating collaboration, integrating multiple fact-checking tools, using analytics and ultimately through time savings and increased efficiency and effectiveness. The EDMO fact-checking platform is also open to journalists because it is essential that such resources are also available to media organisations. A common platform makes it easier for everyone to cope with the increased scope and complexity of disinformation online, all for the purpose of public interest and providing adequate information to citizens.

EDMO's Truly Media platform offers:

- → a digital collaborative environment where members of the same or different organisations can connect, communicate, and work together on joint fact-checking activities/collaborative investigations;
- → tools to discover and collect content from various online sources (web, social media, online platforms) as well as tools to organise, sort and archive the collected content;
- → tools that support the analysis, verification, and fact-checking of multimedia content (text, images, videos);
- → access to data repositories (e.g., Eurostat, OECD repositories);
- → the platform is modular, allowing linking and API-based integration.

The Truly Media platform was co-developed by ATC — Athens Technology Centre and Deutsche Welle - an international German public media service. Access to the platform is free and fact-checkers retain editorial independence, but in order to gain community access they must meet certain criteria:

> **Focus on the European Union:** Any fact-checking organisation willing to become a user of the platform has to be established in the EU and have a demonstrable focus on the EU and/or its Member States in all, or at least in a significant part of its activities. Being a member of the EU national hubs on disinformation will be considered a sufficient demonstration of compliance with this criterion.

---

57 https://edmo.eu/edmo-hubs/

58 https://edmo.eu/admission-criteria/

Competence: Applicants will have to demonstrate competence in fact-checking; they have to be active projects, having published at least 15 fact-checking articles in the three months before the application. Being signatories of the Code of Principles of the International Fact-checking Network will be considered a sufficient demonstration of compliance with these criteria.

Transparency: Applicants will have to disclose and avoid any potential conflict of interest, including work, consulting activities, share-owning or funding from any company or organisation in the social media / digital media sphere. Applicants must disclose their organisational and proprietary structure, and be free of influence or control of political parties or movements.

Ethics: Applicants must comply with the applicable rules of ethics in their area of expertise.

EDMO clearly states both the entire procedure and duration of assessing and approving applications of fact-checking organisations wishing to participate in the EDMO network. Upon receipt, applications are sent to the External Assessment Committee, composed of ten experts chosen by the EDMO Executive and Advisory Boards. Within four working days the Committee makes a recommendation as to whether the application should be accepted or not. The Committee's recommendation is confirmed by the members of the EDMO's Advisory Board, as additional validation and verification. If the application has been rejected, the candidate has the opportunity to appeal once within seven calendar days, with additional information addressing the rejection criteria.

EDMO currently brings together 31 European fact-checking organisations and projects. In addition to specialised fact-checking organisations, it also includes media and journalistic projects such as: international news agency Agence France-Presse (AFP), German news agency Agency DPA, whose shareholders are 174 German media outlets[59], the Baltic Center for Investigative Journalism Re:Baltica, a non-profit organisation dealing with investigative journalism in the public interest. Another example is Science Feedback, a non-profit organisation that verifies the credibility of influential claims and media reporting claimed to be scientific, primarily in the field of climate and health.

Fact-checking organisations that are part of the EDMO network are expected to contribute with their expertise and experience to revealing false narratives, especially those that have a cross-border dimension.

Complementary to EDMO activities, in May 2021 the European Commission issued a call[60] to draft a Code of Professional Integrity for Independent European Fact-checking Organisations and OSINT (Open Source Intelligence)[61], and to manage its implementation and support organisations in the process of compliance with the standards of the Code. A consortium of six European organisations dealing with fact-checking and disinformation research leads the process of creating the Code: AFP (France), Correctiv (Germany), Demagog (Poland), Pagella Politica/Facta (Italy) and EU DisinfoLab (Belgium), led by Fundación Maldita.es (Spain). The consortium operates under the title The European Fact-Checking Standards Network Project[62], and the Code should come to life during 2022 and 2023. The key standards set out in the description of the objectives include independence, transparency and methodological and journalistic quality of action[63].

---

59 **https://www.dpa.com/en/company#company**

60 **https://digital-strategy.ec.europa.eu/en/funding/call-proposals-integrity-social-media**

61 Organisations using techniques and tools to collect public information, correlate data and process them in order to gain useful and applicable knowledge in various fields.

62 **https://eufactcheckingproject.com**

63 **https://www.disinfo.eu/projects/european-fact-checking-standards-project/**

5.6. Council of Europe standards

One of the most recent Council of Europe Recommendations (CM/Rec (2022) 4), adopted by the Committee of Ministers on 17 March 2022[64], contains several recommendations to Member States on promoting a favourable environment for quality journalism in the digital age. The document describes a new information ecosystem which has radically transformed news consumption habits and other media content consumption habits. It points out that there is abundant information online, often accessed through algorithm-driven platforms that lack editorial control and transparency. At the same time, evidence suggests that the business models of online platforms and other intermediaries, which have become a main source of news and information for large audiences, facilitate or even incentivise the spread of sensationalist, misleading and unreliable media content. Information overload impacts people's focus and attention spans and has made it markedly more difficult for many to identify and access quality journalism.

The Council of Europe recommendation, as well as earlier European Commission documents, recognise disinformation and online manipulation as some of the key challenges for democracy. At the same time, in such an environment, political parties and unscrupulous politicians use the term "fake news" to discredit the media and undermine the legitimacy of journalism. Trust in the media, as well as trust in politics, institutions and expertise, has declined to worryingly low levels in many states. Several media outlets that have traditionally been committed to public interest now find themselves unable to maintain their reader or viewer base and struggle to adapt their operations to a digital environment. All this is stated in the document and taken into account by the Council of Europe as context which requires stronger engagement of different actors, especially Member States, in promoting a favourable environment for quality journalism. The recommendations include, inter alia, investing in the quality of journalism, strengthening the role of journalism in the process of fact-checking and credible reporting, and ensuring the financial sustainability of journalism.

The document, for example, states that professional journalism and fact-checking, audience engagement, transparency and greater accountability within media organisations and online intermediaries can contribute to (re)establishing trust and healthy relationships between media actors and the public. It also highlights the need to develop and consistently apply tools, techniques and ethical guidelines that can be applied in the news production process, including reporting of news collected from social media, using user-generated content during emergencies, or using eyewitness videos as evidence. These practices should systematically be integrated into basic journalism training to reinforce fact-checking and careful selection of sources as a cornerstone of quality journalism. In particular, media should exercise vigilance and verify stories originating from anonymous private fora, messaging apps or social media before citing, incorporating or otherwise relaying them, in order to avoid spreading disinformation. Joint fact-checking projects between multiple newsrooms, universities, non-governmental organisations and online platforms, as well as between organisations in different States, can have beneficial effects, especially in pre-electoral and referendum periods, as the Recommendation of the Council of Europe states.

64 https://search.coe.int/cm/pages/result_details.aspx?objectid=0900001680a5ddd0

6.
Croatia: media, trust and disinformation

The contemporary media system in Croatia is decisively shaped by a turbulent historical, political and social framework, as well as by global changes in the media and communication environment.

The directions of development of the media system in Croatia are consistent with the development of media systems in other countries of Central and Eastern Europe (CEE). Peruško et al. (2021: 25-32) argue that media transformation in Central and Eastern Europe is fundamentally determined by four "critical junctures" — modernisation, socialism, post-socialist democratisation and communication juncture. The transition from socialism to democracy, which is the subject of most research papers on media systems in Central and Eastern Europe (cf. Dobek-Ostrowska 2015, 2019; Hallin and Mancini 2013; Herrero et al. 2017; Peruško 2013, 2016; Peruško et al., 2021), as well as changes in the global communication environment, such as the development of new media technologies, digitalisation, mediatisation (cf. Lundby 2014), hybridity (cf. Chadwick 2017) and the like, have markedly determined the development, structure and quality of the contemporary media scene in Croatia.

In addition to temporal stages and processes that Croatia shares with the rest of the CEE, the development of the contemporary media scene in Croatia was strongly influenced by the Homeland War (1991 - 1995). During and after the war, the media were heavily influenced by the state, and on the economic level, the media transition of the 1990s was marked by "nationalization, privatization and marketization" (Peruško 2013: 714). Events from the 1990s paved the way for the development of the contemporary media system that Peruško (2013: 714) refers to as the "polarized pluralist media system" characterised by "the late development of the mass press, weak professionalization of journalism, strong political parallelism, and the strong role of the state" (721). Since the 1990s, the media have played an important role in the democratisation of Croatian society and still remain an important pillar of democracy and a social corrective element.

6.1. Croatian media system

According to Peruško and others (2021: 1), "the transformations of the media following the fall of socialism involved changes in the institutions - in the rules and values, as well as in the practices of media professionals (including journalists) and in media-related practices of audiences." These changes in Croatia were in line with the political and social consolidation of post-socialist society. Bilić (2012: 828-829) identifies three transformative periods in the development of the Croatian media system. The first period lasted from 1990 to 1999 and was marked by the construction of the state and the monopoly of national state television (Croatian Television, HTV). The liberalisation of the media market only referred to the press. The impact of new technologies was marginal. According to Bilić (2012), the second period lasted from 2000 to 2003, and was marked by structural political transformation and democratic consolidation. This was followed by the liberalisation of the television and telecommunications market and the beginning of the transformation of the Croatian Radiotelevision into a public service. The first commercial television with national concession, Nova TV, began broadcasting in 2000, and the second, RTL Television, in 2004. Finally, the third period began in 2004 and was marked, as suggested by Bilić (2012: 827), by aligning national media legislation with the provisions of the European Union due to the Croatian accession to the EU, the increasing importance of the internet and the intensive penetration of "consumer and global pop-cultural products" into the Croatian media market.

In the classification of the Croatian media system, Peruško (2013) relies on the well-known conceptualisation of media systems by Daniel Hallin and Paolo Mancini (2004, 2012) and places Croatia in a "polarized pluralist Mediterranean model". However, several authors adapted Hallin and Mancini's conceptualisation to the specific Central and Eastern European context and gave, probably, an even more accurate assessment of the Croatian media system. Thus, Castro Herrero et al. (2017) use four indicators – political parallelism, the role of public service, press freedom and foreign ownership — to group countries into three clusters: eastern, central and northern (4808). They position Croatia in the central cluster, located somewhere in the middle in terms of political parallelism and media freedom, but also point to Croatia's lower results than the eastern and northern clusters in the category of foreign ownership of media and higher results in terms of the relevance of public service (p. 4810).

Furthermore, Dobek Ostrowska (2015) identifies four models of media and politics in Central and Eastern Europe: "the hybrid liberal", "the politicized media", "the media in transition" and "the authoritarian model" (25). According to Dobek Ostrowska, Croatia belongs to the model of "the politicized media" characterised by "low democratic standards and political culture of societies, a high politicization of public broadcasting service and control over public radio and television by political actors" (2015: 28). The adaptation of Hallin and Mancini's model offered by Peruško (2016) includes four key indicators — "role of the state," "media market," "political and economic parallelism" and "professionalization of journalism" (258).

Peruško applies these indicators to generate three models of media systems involving the countries of Central and Eastern Europe, as well as the countries of Western Europe: the "South/East European model", the "European mainstream model" and the "Scandinavian model" (p. 258). Croatia demonstrates the features of the "South/East Europe" characterised by: "lower to medium quality of public television", "lower newspaper circulation", "higher party influence", "higher owner influence" and "lower professionalism and independence" of journalism (Peruško, 2016: 258). Although the categorisation of countries into proposed media models is subject to change, as these models are based on dynamic categories such as freedom of the press, share of foreign ownership in the media, market volatility, etc., they continue to provide a solid framework for understanding the modern media market in Croatia.

6.2. Croatian media market

The Croatian media market can generally be seen as diverse and competitive, especially given the size of the market and its economic strength, or its absence. The owners of many print media and major commercial television networks in Croatia are foreign corporations. Austria's Styria Media Group owns two of the three best-selling newspapers: *24sata* tabloid (34% weekly reach) and *Večernji list* (16% weekly reach) (RIDNR, 2022: 71). Along with Styria, Hanza Media has the largest share of the print media market. The leading Hanza publication is *Jutarnji list*, with a weekly reach of 23% (RIDNR, 2021). It is the second best-selling publication after 24sata. Styria and Hanza Media also have a leading role in the digital media market. The online edition of *24sata*, 24sata.hr is the second most successful online brand, with a weekly reach of 49% (RIDNR, 2022). The online edition of Hanza's *Jutarnji list*, jutarnji.hr is the third most popular online media brand with a weekly reach of 39% (RIDNR 2022). At the top of the ranking of the most popular online media is Index.hr (weekly reach of 56%, RIDNR, 2022).

As for the radio and television market, Croatian Television (HTV) has been the main source of information for Croatian citizens for years, but in the last ten years this has no longer been the case, because Nova TV's news programme has become the most watched.

Nova TV was owned by Central European Media Enterprise (CME Group) until 2017 and today it is part of United Group owned by the British investment fund BC Partners. The German RTL Group owned seven channels in Croatia: RTL (general), RTL 2 (entertainment), RTL Kockica (children's) and several specialised channels of cable television. In February 2022, it was announced that RTL was sold to a subsidiary company of the Czech investment group PPF - Central European Media Enterprises (CME), the former owner of Nova TV. The television news market is also enriched with cable news channels. Thus, N1 started operating in 2014 as a regional television project and information partner of CNN. It broadcasts via cable and OTT platform, with special content for Croatia. In Sarajevo, Al Jazeera Balkans was founded in 2011 (as part of the Al Jazeera media network), which broadcasts for Croatia, and a part of its programme is re-broadcast on Sports Television (SPTV), a national commercial television, with a small share in viewership.

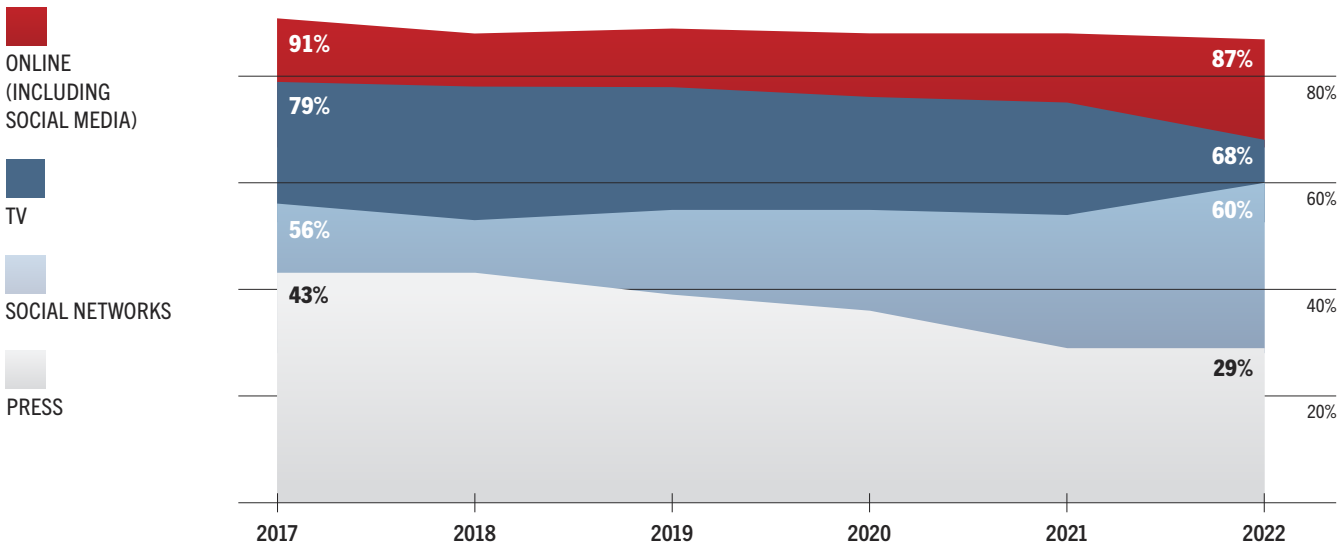
According to the Agency for Electronic Media and the Register of Radio Service Providers, 139 radio broadcasters are currently registered in Croatia, with a total of 154 radio channels. There are also 11 HRT radio channels (3 national and 8 regional) and 15 non-profit radio stations.[65] One of the major problems of the radio system in Croatia is that regional and local self-government units often have a part in the ownership structure of radio stations. This is the case with as many as 64 radio stations, with 26 of them being 100 percent owned by regional or local self-government units. Such ownership constellation suggests that in such cases it is difficult to count on journalistic and editorial impartiality. On the other hand, it raises the question of economic sustainability of local radios in Croatia.

According to a Reuters survey (2022), commercial radio stations with the highest weekly reach are Otvoreni radio and bravo! radio (previously Narodni radio). Both have a national concession and are focused almost exclusively on music (the former on foreign music and the latter on domestic music) as well as casual entertainment. In general, there is little research on the radio market in Croatia, and listening ratings are not widely available and are often subject to selective interpretation.

65 https://www.aem.hr/radijski-nakladnici/

Changes in the national and global media market, which went hand in hand with intense technological changes, fundamentally changed the media habits of the Croatian audience. For example, in 2010, television was the main source of news for most Croatian citizens, while the Internet was the third in importance, but also the source most trusted by citizens (cf. Brautović, 2010). Today the situation is fundamentally different. According to the RIDNR (2022), online media (including social networks) are the source of news for as many as 87% of respondents in Croatia (see Graph 5). Television is ranked second; it is a source of news for 68% of respondents, which is a significant decrease compared to 2021 when this figure was 76%. Social networks are the most important source of news for 60% of respondents, which can be viewed as a surge if compared to 2021 when the figure was 54%. The print media is the least represented source of information in Croatia (for 29% of respondents this year as well as last year).

Sources of News 2017- 2022



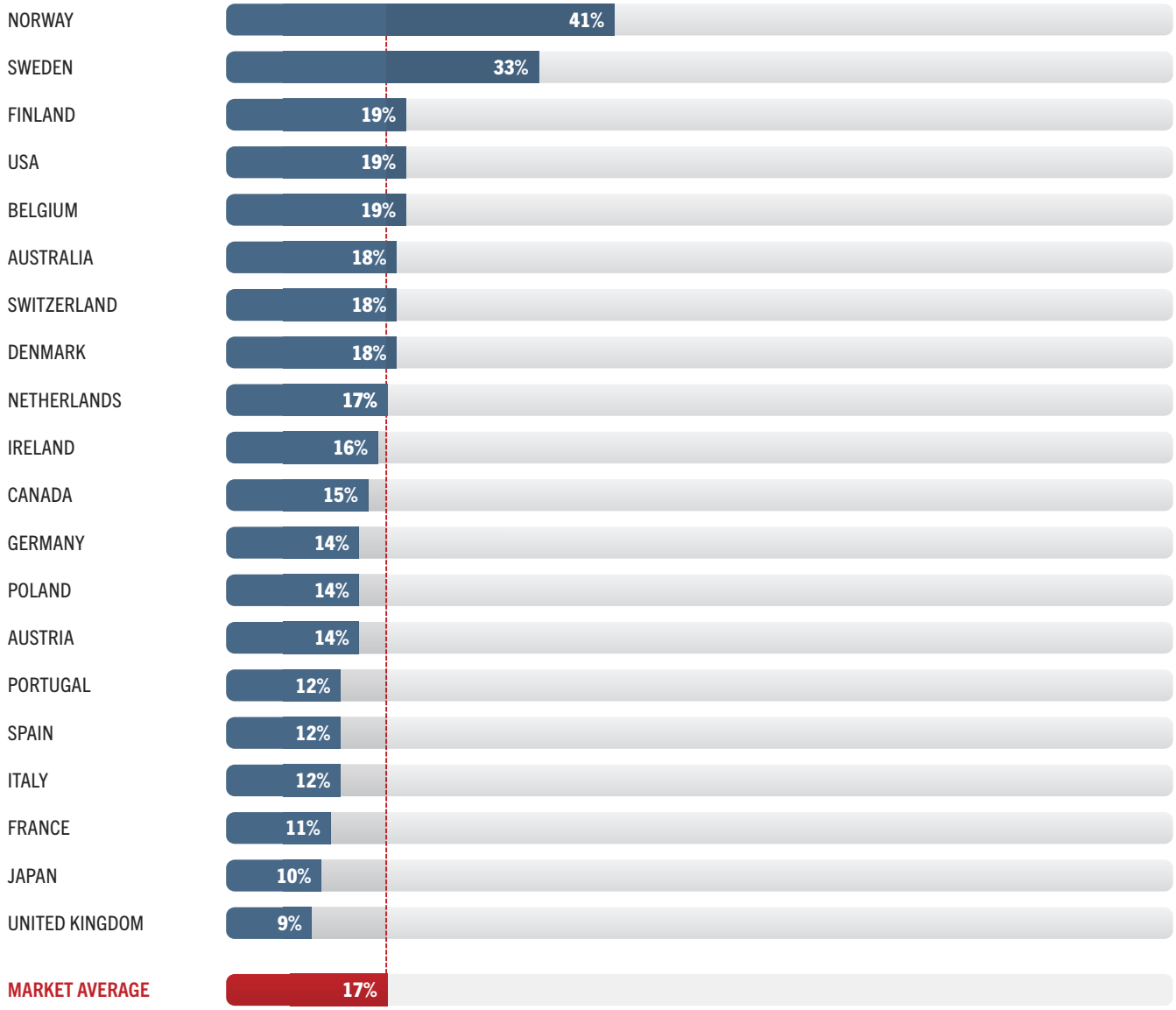
Graph 5: Sources of news in Croatia. **Source:** RIDNR, 2022

The most popular social network in Croatia is Facebook, with 57% of those who use it as a source of news and 74% of those who use it as a source of news and for other purposes. It is followed by YouTube, which is used as a source of news by 30% of respondents and by 71% of respondents who use it for news and other content (RIDNR, 2022: 71). It is interesting to notice that compared to last year in Croatia, the intention of the audience to search for news on YouTube, where unverified and manipulative content is quite common, is on the rise.

Changing media habits have imposed new challenges to publishers and encouraged the implementation of new business models, such as the introduction of various forms of paywall for digital content. Digital editions of two leading political daily newspapers, *Jutarnji list* and *Večernji list*, have recently presented a variant of "soft" paywall (cf. Nenadić and Ostling, 2018), which offers free access to a limited number of articles or charges for "premium" content. Graph 6 shows the trends in the market of paid-for online news. It is interesting to notice that among the markets included in the Reuters Institute research, the highest percentage of those willing to pay for access to online news is recorded in Norway and Sweden, markets that otherwise stand out by investing in quality journalism and a high rate of trust in the media.

Proportion who paid for online news in last year (subscription, membership, donation, or one-off payment) – selected markets

Have you paid for ONLINE news content, or accessed a paid-for ONLINE news service in the last year?
(This could be digital subscription, combined digital/print subs or one-off payment for an article or app or e-edition).



Graph 6: Payment for online news. **Source:** RIDNR, 2022

Changes in the media market and in the media environment have caused disruptions in advertising revenue. According to data from the Croatian Association of Communications Agencies (HURA, 2021), advertising expenditure is increasingly shifting from traditional to digital media, although television remains a medium outlet with the highest advertising revenues. Compared to 2019, 2021 saw a significant decrease in advertising revenues in the press media (-29%) and radio (-15%), while there was an increase in revenues from advertising on television (1%), especially on the Internet, not including social networks (14%)[66].

66 https://hura.hr/istrazivanja/medijska-potrosnja-u-hr/

6.3. Trust in media and institutions

Changes in the media economy and media consumption are not necessarily compatible with trust in the media. Trust in the media in Croatia is among the lowest in the European Union and if judged by Eurobarometer surveys, it has a downward trend. As can be seen from Table 1, trust in the media in Croatia amounted to only 28% in spring 2021.

Public opinion in the European Union

How much trust do you have in certain institutions? For each of the following institutions, do you tend to trust it or tend not to trust it?

The media (%)

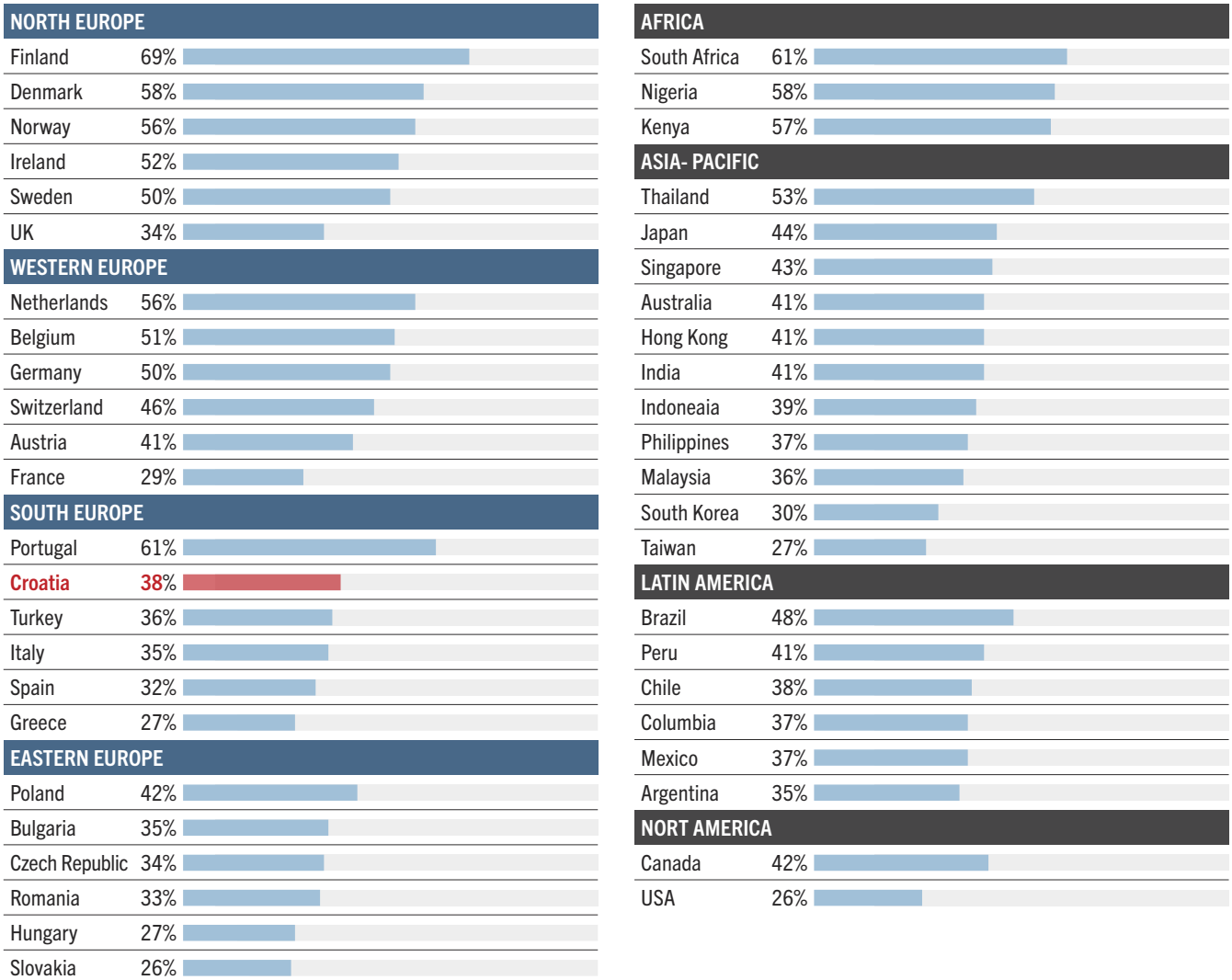
	Tend to trust		Tend not to to trust		I don't know
	Sp.2021	Diff. Sp.2021 Sum.2020	Sp.2021	Diff. Sp.2021 Sum.2020	Sp.2021
EU27	41	1	56	-1	3
EURO AREA	40	1	57	-1	3
NON-EURO AREA	43	1	53	-1	4
Belgium	47	1	53	0	0
Bulgaria	39	0	51	-1	10
Czech Republic	49	-2	49	3	2
Denmark	57	-9	40	9	3
Germany	46	-4	51	4	3
Estonia	52	0	48	0	0
Ireland	53	2	47	-2	0
Greece	18	-3	81	3	1
Spain	31	-2	66	1	3
France	26	5	71	-5	3
Croatia	**28**	**-9**	**68**	7	**4**
Italy	40	5	56	-4	4
Cyprus	28	-10	70	10	2
Latvia	41	-1	58	3	1
Lithuania	40	-3	60	5	0
Luxembourg	40	-7	56	3	4
Hungary	34	-1	65	3	1
Malta	25	4	68	0	7
Netherlands	59	3	40	-3	1
Austria	49	-13	46	11	5
Poland	42	3	53	-3	5
Portugal	62	0	36	-1	2
Romania	44	7	52	-8	4
Slovenia	37	4	62	-3	1
Slovakia	36	-3	61	1	3
Finland	75	4	25	-4	0
Sweden	53	-4	47	5	0

Table 1: Trust in the media in the EU. **Source:** Standard Eurobarometer 95, spring 2021

The causes of such low trust in 2021 should be found, among other things, in events related to the COVID-19 pandemic. A Reuters report from 2022 investigating whether respondents believe "most news most of the time" suggests that the trust rate in news in Croatia is 38% (see Graph 7), which represents a decrease compared to 2021 when this figure was 45%.

Proportion that trusts most news most of the time – all markets

Thinking about news in general, do you agree or disagree with the following statements? – I think you can trust most news most of the time.



Graph 7: Trust in the news. **Source:** RIDNR, 2022

Eurobarometer surveys which measure trust in certain types of media, suggest another indicative figure. **In fact, trust in traditional media in Croatia is below the average of the European Enion, while trust in social networks is above the EU average.** Table 2 shows comparative data for 2019 and 2021/2022 (Standard Eurobarometer 92, Autumn 2019 and Standard Eurobarometer 96, Winter 2021 /2022).

	2019		2021/2022	
	CROATIA	EUROPEAN UNION	CROATIA	EUROPEAN UNION
Radio	50%	57%	46%	56%
Television	47%	49%	43%	49%
Press	39%	46%	36%	49%
Internet	39%	32%	32%	35%
Social networks	32%	20%	27%	20%

Table 2: Trust in different types of media

For many years, radio has been the most trusted media outlet in Croatia, but also at the level of the entire European Union. At the level of the whole EU, television and the press are the most trusted media outlets after radio, followed by the Internet and social networks as the least trusted. In Croatia, television is the most trusted media outlet after radio, the press and the Internet, and the least trusted are social networks.

Asked in the last wave of the Eurobarometer survey (winter 2021-2022) whether they agreed with the statement that information about political events published on social networks could not be trusted, 64% of respondents from Croatia agreed and 29% disagreed.

The most reliable media brands according to RIDNR (2022) are two commercial television networks, Nova TV and RTL, which are trusted by 58% of respondents according to the report. Although these two commercial television networks are the most trusted media organisations, they both record a big fall in trust which in 2021 amounted to 75% for Nova TV and 73% for RTL. The decline in trust is also noted by other brands. According to the same source, in 2021 59% of respondents trusted public television, HTV, compared to only 43% in 2022. In 2022, HTV also recorded the highest rate of distrust of all the media included in the Reuters survey (32%). HTV is followed by the most read Croatian news portal Index.hr, which is not trusted by 30% of respondents.

As for institutional trust, it should be noted that Croatia is a country with continuously low levels of trust in political institutions, one of the lowest in the EU (Bovan et al., 2022; Eurobarometer 2021, 2022; Henjak 2017). According to a Eurobarometer survey in autumn 2021, 77% of Croatian respondents said they did not trust the Croatian government; 22% of respondents trusted the Croatian government, which is significantly less than the average of trust in national governments at the level of the whole European Union, which was then 37%. According to the latest survey from winter 2021 /2022, the percentage of trust in the Croatian government was 21%, while the EU average was 35%.

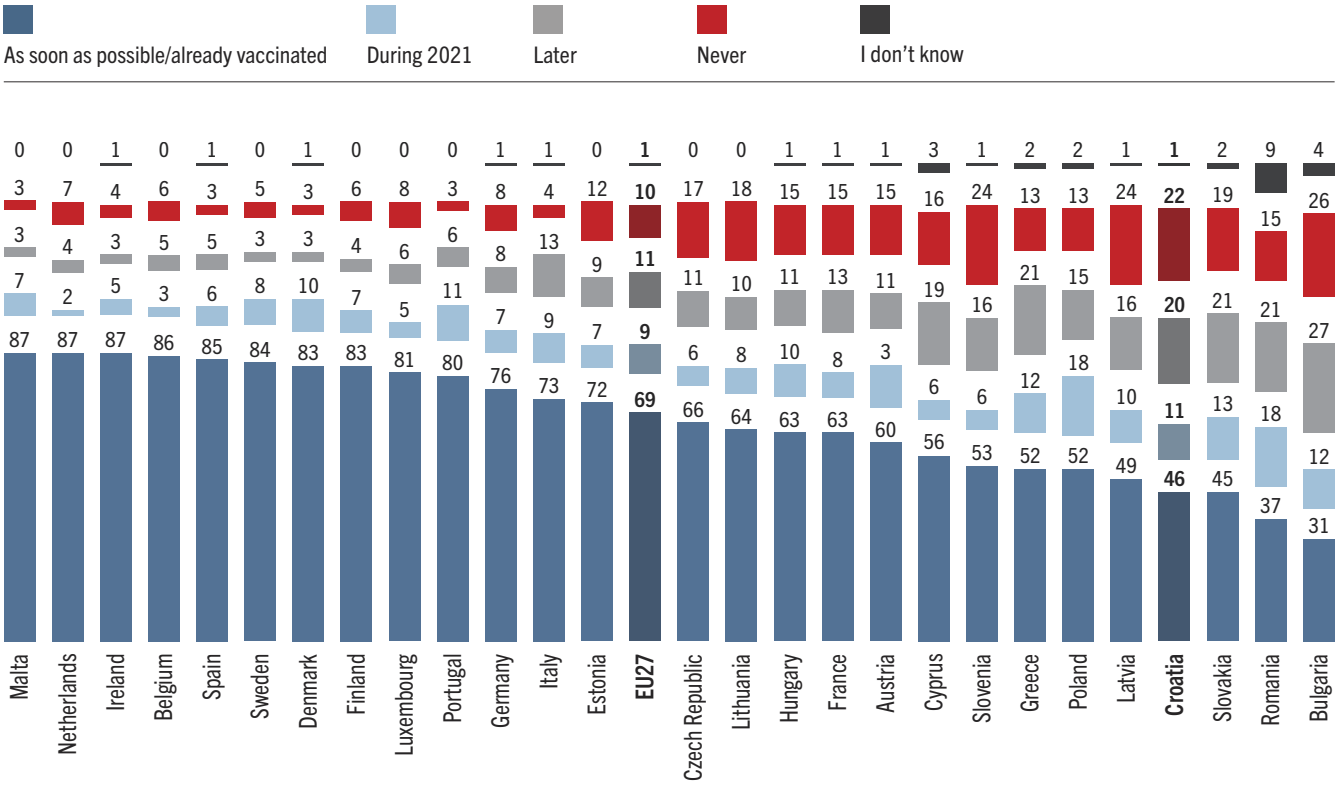
The level of trust in the Croatian Parliament in 2021 was also low (22%), while the trust in the judiciary was even lower according to the same survey; in 2021 it was 21%, compared to 54% at the level of the whole European Union. According to the latest survey from winter 2021 /2022, trust in the Croatian Parliament was 22%, while the average trust in national legislative institutions at the EU level amounted to 36%. At the beginning of this year, trust in the judiciary was at 24% in Croatia and at the Union level it was at a stable 54%.

6.4. Trust in institutions during the COVID-19 pandemic

It is to be assumed that the context of institutional distrust has contributed to distrust in the dominant official narrative and policies related to the COVID-19 pandemic, which has prepared grounds for disinformation, misinformation and "alternative" narratives to spread.

Several figures and research studies corroborate this. First, vaccination rates in Croatia are among the lowest in the European Union. According to official data, they were 68.68% in April 2022[67]. By way of comparison, in Ireland, France, Denmark or Germany this figure is more than 90%. According to a Eurobarometer survey from spring 2021, 22% of respondents from Croatia said they never wanted to get vaccinated against COVID-19, which is one of the highest percentages in the EU (see Graph 8).

QA21 When would you like to get vaccinated against COVID-19 (coronavirus)?
(%)



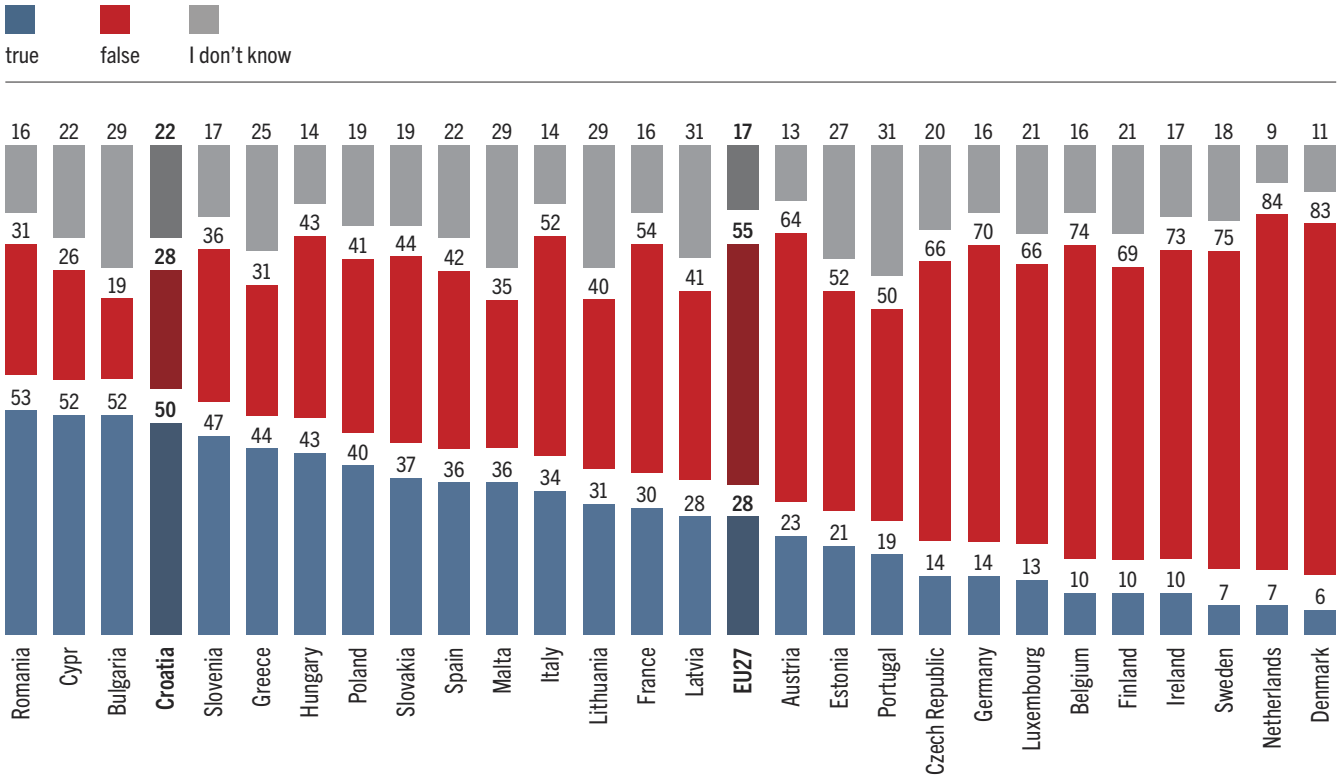
Graph 8: Attitude toward vaccination in EU countries. **Source:** Eurobarometar 95, spring 2021

According to the winter 2021/2022 Eurobarometer survey, this percentage in Croatia was 17%, while the European Union average was only 8%. Secondly, it is quite concerning that the level of trust in healthcare professionals recorded in Croatia in two consecutive rounds of Eurobarometer measurements, in spring 2021 and winter 2021/2022, was low. In both cycles, 66% of Croatian respondents said they trusted healthcare professionals, while 32% did not trust them, which is a decrease of several percent compared to earlier measurements. The average EU-wide trust in healthcare professionals is around 80%.

67 https://www.koronavirus.hr/560-novih-slucajeva-u-protekla-24-sata-utroseno-520-doza-cjepiva/35

Third, some studies suggest that a certain percentage of Croatian citizens are prone to conspiracy theories. For example, the Special Eurobarometer 516: European citizens' knowledge and attitudes towards science and technology from September 2021[68] suggests that as many as 50% of Croatian citizens believe that viruses have been produced in government laboratories to control our freedom (see Graph 9).

For each of the following statements, please indicate whether you believe them to be true or false, or you do not know:
Viruses have been produced in government laboratories to control our freedom (%)



Graph 9: Believing the claim that viruses have been produced in government laboratories. **Source:** Special Eurobarometer 516

The propensity to conspiracy theories related to COVID-19 in Croatia was investigated by Mirjana Tonković et al. (2021). The authors found that nearly a quarter of respondents agree with conspiracy theories related to COVID-19 and that belief in corona-related conspiracies is associated with lower levels of education, lower economic standard, greater importance of religion and lack of political self-identification of respondents (2021: 7). Likewise, the authors showed that the strongest predictor of belief in conspiracy theories is distrust in science. In addition to this study, several recent scientific studies also analysed trust in institutions in Croatia during the pandemic.

Building on research papers that have found that institutional trust is positively linked to the acceptance of pandemic measures (see e.g., Han et al., 2021; Wingen and Schreiber, 2020), Kosta Bovan et al. (2022) found that people who were satisfied with the public institutions managing the pandemic crisis had a higher degree of institutional trust, as well as those who felt that Croatian

68 https://europa.eu/eurobarometer/surveys/detail/2237

citizens had a "shared identity" (12-13). On the other hand, the authors determined that belief in conspiracy theories was associated with low institutional trust. They conclude that in uncertain times, in order to strengthen institutional trust, a well-developed professional strategy should be designed to combat fake news and conspiracy beliefs, especially on social networks (2022: 13).

Similar findings were also reported by Dragan Bagić and Adrijana Šuljok (2021) who concluded that trust in state institutions managing the crisis was one of the key factors in accepting restrictive and protective pandemic measures, with the perception of risk as the most important factor (142). Bagić and Šuljok found that among all institutions involved in combating the pandemic for which confidence was measured, respondents had the least degree of trust related to the objectivity of media reporting (see Table 3).

	I don't trust it at all	I mostly don't trust it	I mostly trust it	I have a lot of trust	No response
Trust in National Civil Protection Headquarters and its management of crisis caused by the coronavirus epidemic	12%	21%	34%	31%	2%
Trust in scientists-researchers' ability to find a cure or vaccine for the disease caused by the coronavirus	5%	9%	38%	45%	3%
Trust in the media regarding objective reporting on the crisis caused by the coronavirus epidemic	20%	29%	34%	15%	2%
Trust in the Government of the Republic of Croatia regarding its management of the crisis caused by the coronavirus epidemic	20%	25%	31%	23%	1%
Trust in the health system of the Republic of Croatia in terms of its dealing with the crisis caused by the coronavirus epidemic	6%	15%	40%	38%	2%

Table 3: Trust in the institutions and systems involved in combating the pandemic. **Source:** Bagić and Šuljok, 2021: 134

So, what does all this data tell us about the media environment in Croatia and the potential for disseminating disinformation? First, the transition from traditional to digital media as the main source of news is consistent with changes in audience behaviour in other European countries (RIDNR, 2020: 60-85; RIDNR, 2021: 61 - 108; RIDNR, 2022: 61-109). Second, trust in the media and traditional media is below the EU average. This can be attributed, at least in part, to a general distrust in institutions in Croatia.

While trust in radio and television is still bigger than trust in the internet and social networks, trust in information obtained on social media is above the EU average. In addition, the percentage of people for whom social networks are the primary source of information in Croatia is growing, while television as the primary source of news has seen a significant decline. The most trusted brands are two commercial television networks, Nova TV and RTL but the percentage of people who trust these brands, if judged according to a survey by the Reuters Institute, has fallen substantially. Likewise, print media are continuously losing their audience. This position of newspapers and magazines can be explained by the progressive tabloidisation of their content, but even more so by a digital turn in the media industry that has prompted audiences to stop buying newspapers and look for news online,

where it is still available mostly free of charge (RIDNR. 2020: 10). Publishers have been trying to compensate for the loss of newspaper readers by introducing paywall in their digital editions. Third, the greatest distrust of the Croatian audience this year was expressed towards Croatian television public service, HTV. Finally, research related to the COVID-19 pandemic indicates that a proportion of citizens do not believe that the media provide credible and objective information related to the pandemic, and that a significant number of citizens are prone to conspiracy thinking. All this opens up space for the spread of disinformation, misinformation and other manipulative content.

A Eurobarometer survey conducted in autumn 2019 found that 73% of respondents from Croatia felt they were often exposed to disinformation and fake news. Two years later, according to the Eurobarometer from winter 2021/2022, this figure climbed to 80%. In 2019, 81% of respondents considered this was a problem in Croatia, and today 85% of respondents think so. When asked whether they could recognise disinformation and fake news, in 2019, 70% of respondents from Croatia replied that such information could be easily or relatively easily recognised, while this score at the European Union level was significantly lower, 58% (Standard Eurobarometer 92, Autumn 2019). Today, both Croatian citizens and EU citizens are even more confident when it comes to identifying disinformation, so 73% of respondents in Croatia think they can recognise such news, and 62% of respondents from the European Union think the same. The overwhelming majority of citizens believe that disinformation or fake news is a threat to democracy (90% in Croatia and 81% in the EU, Eurobarometer 96, winter 2021/2022).

6.5. Media freedom and pluralism

The three fundamental laws regulating media in Croatia are the Electronic Media Act, the Media Act and the Croatian Radiotelevision Act. Although institutions are nominally committed to promoting the diversity of media content, Croatia has faced challenges since its independence in terms of competitiveness and media pluralism. In this regard, Peruško et al. (2011: 174) argue that in countries where the state has stopped insisting on media diversity and pluralism through the legislative framework, pluralism has diminished under the relentless influence of the media market.

The Agency for Electronic Media is the regulatory body for electronic media in the Republic of Croatia, which, in addition to regulatory affairs, leads key national media literacy projects. Since 2020 it has managed the project promoting journalistic excellence by awarding grants to journalists for papers in electronic publications for the topics of public interest. Similarly, it has managed the Fund for the Promotion of Pluralism and Diversity of Electronic Media. The Fund supports the production and broadcasting of content of public interest, such as programs that promote the rights of national minorities, exercise of citizens' rights to public information, gender equality, environmental protection, education, science, art and other topics prescribed by the Electronic Media Act. The Fund is financed by 3% of Croatian Radio Television revenues generated from license fees, and in accordance with the Croatian Radiotelevision Act. On average, it amounts to about 33 million kuna annually. Media pluralism is further encouraged through state subsidies to minority media and reduced VAT on the daily press (Peruško et al. 2021: 154).

The most detailed assessment of the state of media pluralism in Croatia was carried out by Bilić et al. (2021, 2022) as part of Media Pluralism Monitor (MPM), a project launched to measure the "potential risks to media pluralism" in EU member states and candidate countries. The monitor uses four themes and accompanying indicators to measure the risk: 1) "fundamental protection" (protection of freedom of expression and the right to information; journalistic profession, standards and protection;

independence and effectiveness of the media authority; universal reach of traditional media and access to the internet); 2) "market plurality" (transparency of media ownership; news media concentration; online platforms concentration and competition enforcement; commercial and owner influence over editorial content; media viability); 3) "political independence" (political independence of media; editorial autonomy; audiovisual media and online platforms in elections; state regulation of resources and support to the media sector; independence of the public media service); and 4) "social inclusiveness" (access to media for minorities; local/regional media and community media; women in the media; media literacy and protection against illegal and harmful speech) (p. 4).

The results are presented on a scale from 0 to 100 percent, with scores between 0% and 33% considered low risk, from 34% to 66% medium risk and from 67% to 100% high risk. In 2022, the results for Croatia show medium risk in all four main areas of the MPM: basic protection (43%), market plurality (65%), political independence (63%) and social inclusiveness (62%). There are no significant deviations from the previous year. However, as MPM increasingly measures the market situation for media and journalism and introduces new digital indicators, the poor quality of market data, outdated data and lack of data, especially in the digital environment (Bilić et al., 2022) constitute a major problem for measurement.

Market plurality in the online environment has a high risk of 72%, mainly due to the risks associated with the viability of journalism in the online environment. As Bilić and others point out (2022: 15) according to the latest available estimates, Google and Facebook dominate the digital advertising market in Croatia. However, the authors also stress that there is no reliable data on the revenue of originally digital (news) media in 2021, since official statistics do not distinguish originally digital media from online editions of traditional media and online platforms. The Association of Advertisers records total online advertising figures, without distinguishing news media and other types of media, while the regulator records figures on the revenue of electronic media, without distinguishing originally digital media from online editions of traditional media. The lack of these key market indicators makes it impossible to adequately assess the situation, and thus makes it difficult to create policies to ensure and strengthen media pluralism. The current situation also points to problems in the field of journalistic standards and protection of journalists, as journalists are particularly threatened by the so-called SLAPP lawsuits. In the field of social inclusiveness, as pointed out by Bilić et al. (2022: 8), access to media for women and media literacy are indicators of high risk. While on public and commercial television women often hold top editorial positions, they are still underrepresented at the executive and management levels of public service and commercial television networks and are often portrayed stereotypically in the news. As Croatia does not have a national policy for media literacy, projects implemented by the Agency for Electronic Media and NGOs, are crucial for the development of media literacy.

In 2022, the World Press Freedom Index placed Croatia in the 48th place out of 180 countries in terms of media freedom. This is an improvement compared to 2021 when Croatia was ranked 56th, and especially compared to 2020 when it was ranked 59th. In the Freedom House report for 2022, Croatia was classified as a "free country" with "highly polarized media" that is "generally free from overt political interference and manipulation". However, journalists continue to face "threats, harassment, and occasional attacks."[69] Despite a satisfactory degree of freedom and journalism autonomy, Reporters Without Borders still point to significant problems in media freedom and the status of journalists, which is in line with the conclusions of the Media Pluralism Monitor.[70]

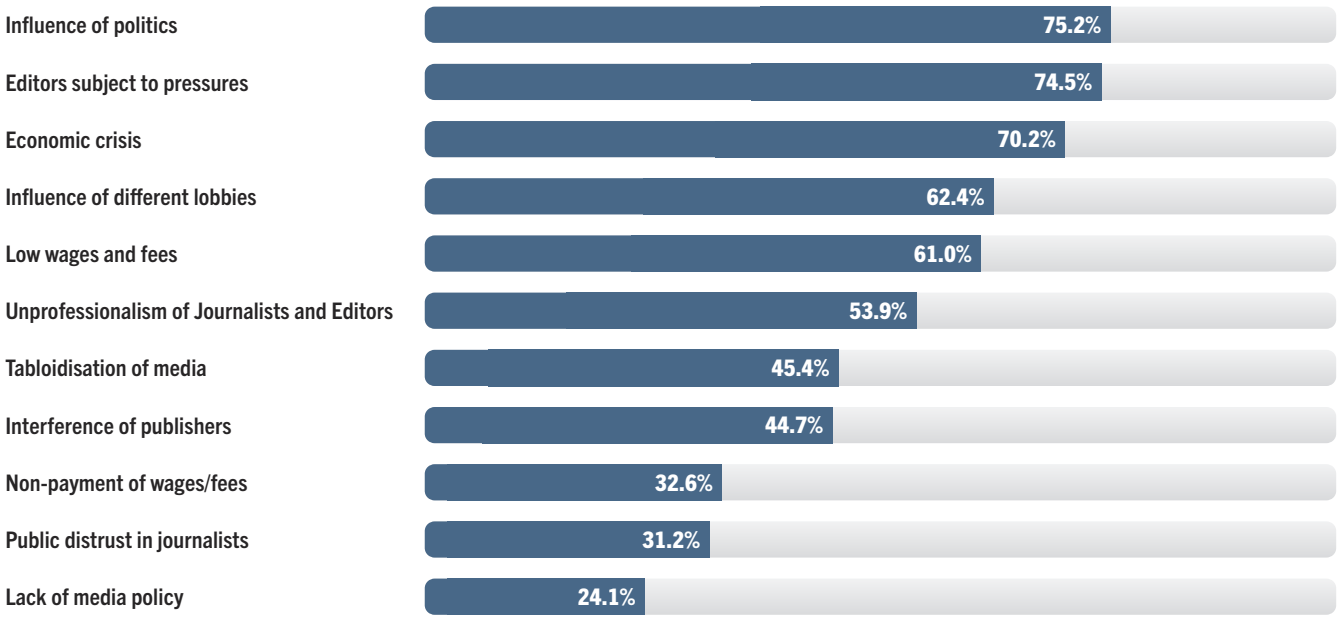
---

69 https://freedomhouse.org/country/croatia/freedom-world/2022

70 https://rsf.org/en/country/croatia

Ivanuš (2021) conducted a survey among 141 journalists and editors of print, electronic and online media in order to examine their perception of professional standards; the pressures they are exposed to; internal and external factors that have an impact on the work of journalists and editors; the correlation between journalists' education, work experience and respect for professional principles; and editorial censorship in Croatian media due to which journalists do not follow professional rules. The majority of respondents, 75.2% of them, believe that political interference is the main problem of the media in Croatia (see Graph 10).

What do you consider the biggest problem of the media in Croatia today?
N=141



Graph 10: Perception of the biggest media problems in Croatia. **Source:** Ivanuš, 2021

Political and commercial pressure on editors was indicated as a problem by 74.5% of respondents, and the influence of various lobbies was indicated by 62.4% of respondents. The economic crisis caused by COVID-19 was recognised as a significant problem by 70.2% of respondents and low professional standards of journalists and editors by 53.9% of respondents. The tabloidisation was highlighted by 45.4% while the owners' influence was mentioned by 44.7% of the editors and journalists involved in the survey. The loss of public trust in the media was identified as a problem by only 31.2% of respondents, which Ivanuš interprets as a sign that journalists and editors have come to accept it (p. 84-85).

The acceptance of loss of trust in the media as something normal is alarming, especially in the context of increasing evidence of the collapse of the quality of media discourse in Croatia.

6.6. "Deprofessionalisation" of journalism

Digitalisation, commercialisation of content and introduction of sensationalism as a new standard in journalism have negatively affected journalistic practice and professional standards of Croatian media (cf. Beck et al., 2021; Hromadžić, 2013; Labaš and Ciboci, 2011; Švob-Đokić et al., 2011; Vrtič and Car, 2016). Grmuša and Prelog (2020) suggest that digitalisation has contributed to the devaluation of professional standards. Croatian journalists working in integrated newsrooms rely heavily on social networks and digital sources to meet the imperative of speed and exclusivity "which often results in the publication of unverified information, which in turn reinforces distrust in the media "(76).

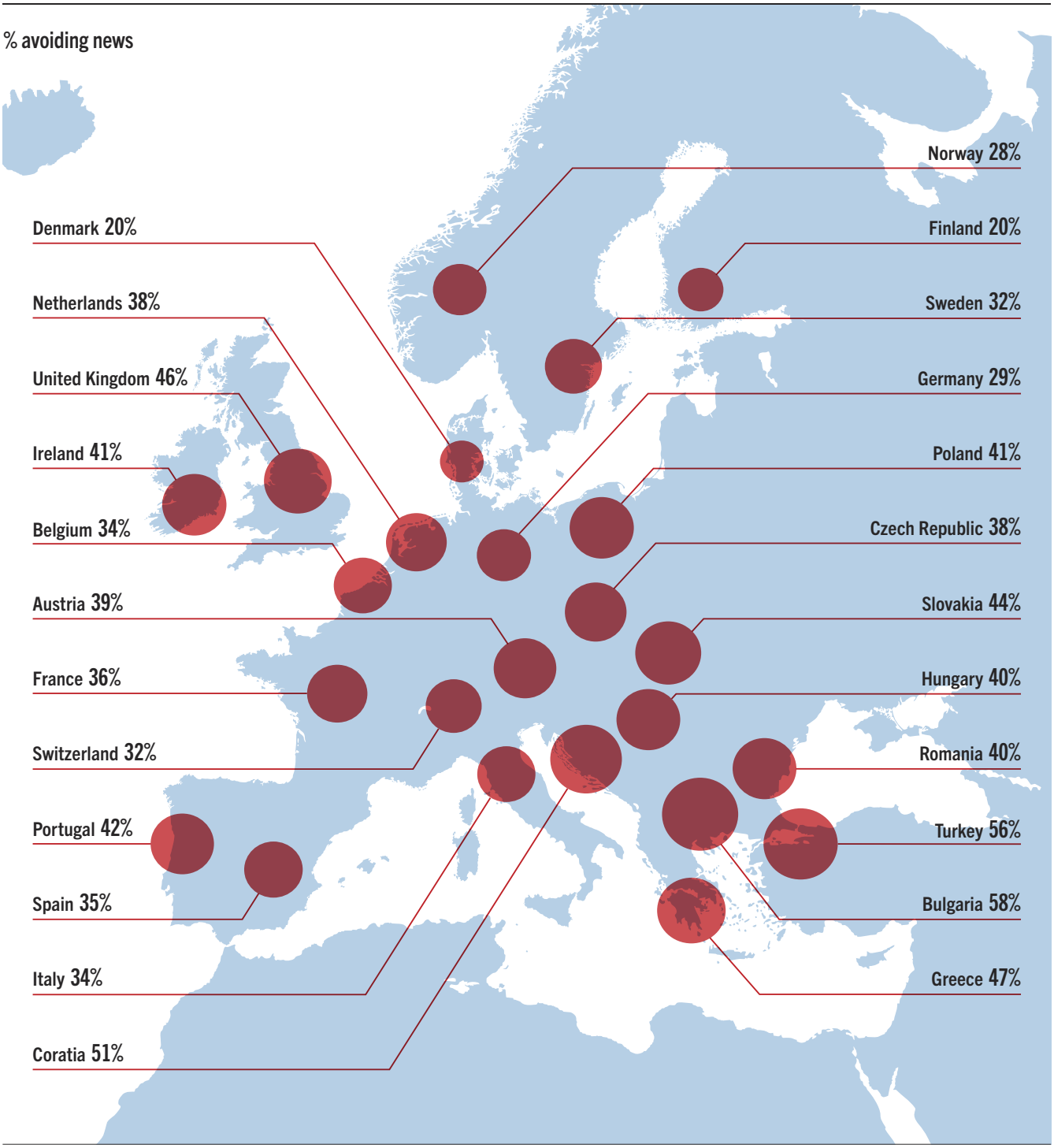
In one of the first studies on the use of social networks as sources of news for traditional media, Volarević and Bebić (2013) found that Facebook and Twitter are emerging as increasingly relevant sources of television news. Vesnić Alujević et al. (2020) found that online media in Croatia in 2015 and 2016 had relied heavily on Facebook as a source of political information.

Today, the media in Croatia, especially online media, almost regularly use social media posts (Facebook, Twitter, Instagram) as sources of news, often by uncritically uploading posts or literally copying them, which again raises the question of journalistic responsibility.

Beck et al. (2021) also show how using clickbaits in headlines, which has become standard practice in Croatian online media, can be manipulated by context. They recognise ten types of clickbaits that use different manipulation techniques to attract readers to click on the content offered: (1) a clickbait that is not untrue, (2) concealing the place of the event, (3) concealing the timing of the event, (4) delusion around event participants, (5) polysemia, (6) anomination, (7) metaphorisation, (8) hyperbolisation, (9) conditional news, and (10) a completely untrue title.

Another important topic related to the digital revolution is the disappearance of the boundary between professional and private use of social networks. In her analysis of journalistic practices on Twitter in Croatia, Nenadić (2020a: 142) establishes that "journalists in Croatia continue to recognise the traditional boundaries of their work into the social media realm (...); they identify themselves through what they do, and, very often, in relation to who their employer is, (...); they mainly communicate with each other and share each other's content, rarely including other (ordinary) users in the process of information exchange and verification. At the same time, journalists are adopting the characteristics of the social media environment, making their communication faster, shorter, more immediate, often more emotional and commentary based, and including a significant level of self-promotion, which represents a deviation from traditional journalistic and media principles." Nenadić proposes the introduction of social media guidelines that would be developed in partnership with journalists. These guidelines would serve "as a means of self-regulation, or as an agreement between journalists and media organisations on what is desirable, and what would be counted as the potentially problematic use of social media" (52). Media outlets in Croatia have generally not yet proposed similar guidelines that would set the framework for professionally acceptable communication on social networks or at least these guidelines are not public. Nothing similar has been proposed either by professional association. The mechanisms of media self-regulation in Croatia are generally very unclear. Ivanuš (2021: 87-88) warns that 44.7% of journalists involved in her research claim that they are not sufficiently familiar with self-regulation (55.3% think they are), 60.3% think they do not have enough information about the importance of self-regulation, and as many as 81.6% of journalists and editors think that self-regulation is important.

All these factors contribute, to some extent, to "media fatigue", which results in active avoidance of news, which amounted to 51% in Croatia this year, which is one of the highest results recorded by the Reuters Institute Digital News Report in Europe (see Graph 11).



Graph 11: News avoidance. **Source:** RIDNR, 2022

In 2019, the Reuters survey revealed that Croatia was the country with the highest percentage of people who "found themselves actively trying to avoid the news" (56%) of all the countries involved in the survey.

Another study by scientists from the Faculty of Political Science in Zagreb, conducted within the framework of the JOURLAB project, shows that people who consciously or unconsciously avoid

news in Croatia, do so because the news is too negative and incites conflicts and tensions in society.[71] Another recent finding also points to the complex relationship between Croatian media and the public.

The Reuters Institute report from 2021 shows that a large majority of Croatian respondents (75%) are not concerned about the financial situation of commercial media organisations in Croatia, while only 15% are concerned and 9% do not know. Likewise, only 28% of respondents believe that the Government should intervene to aid commercial news organisations facing financial difficulties, 47% think that it should not, and 25% do not know.

Negativity, as one of the most prominent features of political news reporting in Croatia (Grbeša and Šalaj, 2018: 167; Vlašić Smrekar and Ivančić Belošević, n.a) is very often embedded in the populist narrative, marked by a Manichaean rift between "ordinary people" and "corrupt elites". The media actively put themselves on the side of the "people" against the establishment, which can be interpreted using Strömback and Esser's (2009) theses on "media interventionism". They argue that media interventionism refers to "a media-centred political reporting style, in which increasingly, journalists and media actors become the stories' main newsmakers rather than politicians or other social actors" (2009: 217). They refer to the famous study by Blumler and Gurevitch (1995) on the crisis of public communication in which Blumler and Gurevitch argue that "journalistic attitudes toward interventionism thrive in political cultures where public opinion is more cynical and distrustful of political institutions" and that a climate is created in which "adversarial journalism seems socially acceptable" (218).

The presence of negativity and media populism in the Croatian media space has been extensively documented by journalists Nataša Vlašić Smrekar and Tanja Ivančić Belošević in their analysis carried out within the project of the Society for the Protection of Journalists' Intellectual Property Rights.[72] The authors in their study conclude that "the Croatian media are not immune to populism and that the degree of its presence depends both on the medium and on the topic in question." Likewise, they found that two-thirds of the articles covered by the analysis are negative towards politicians while sensationalism is present in more than 90 percent of the articles. Among other things, they use the Croatian term "uhljeb"[73] as an example of populist discourse in the Croatian media, which, they conclude, is "so used in certain media that it is starting to lose weight and its true meaning."

71 **http://zagrebnewslab.eu/jourlab/**

72 **https://dznap.hr/medijski-populizam/**

73 "A widely used derogatory term in Croatian public discourse referring to someone who has been employed in the public sector through a nepotistic relationship or political party affiliation, normally without the required skills or qualifications for that position" (Šimić Banović, 2019).

7. Conclusion and recommendations

The global communication environment has profoundly changed with the emergence of new technologies and actors. The development of the Internet and social networks has democratised public space and enabled citizens to participate in a public debate in a way that was previously impossible. However, accelerated digitalisation and platformisation have also brought several challenges for citizens, traditional media and journalism as a profession. The digital transformation of the media and the information sphere has completely changed the "media diet" of audiences who are increasingly turning to online search engines, social networks and non-journalist sources when looking for news. The monopoly of tech corporations over the digital public sphere, algorithmic content management, the dissemination of unverified, false and manipulative information and narratives, and the instrumentalisation of the digital space to promote radical ideas, have led to an "information disorder" that presents one of the key challenges for modern democratic societies.

The role of the media in this process is twofold — they are both victims and part of the problem. With online platforms dominating digital advertising, the media are forced to develop alternative business models to ensure sustainability and maintain relevance. However, in the process, they often follow the imperative of speed at the expense of accuracy. Likewise, editorial decisions are increasingly based on the "attention economy", web analytics and clickbait headlines rather than on the value of news stories and public interest assessment. This process of "deprofessionalisation" of journalism and the inability of the media to respond adequately to technological changes and disinformation challenges, has opened the space and need for specialised fact-checking organisations.

Societies with low levels of trust in institutions, underdeveloped political culture and low professional standards in journalism, such as Croatia, are particularly vulnerable and subject to disinformation campaigns. This was especially evident during the COVID-19 pandemic, which showed that a certain number of Croatian citizens are inclined to believe conspiracy theories, and that the level of trust in mainstream sources of information, such as state institutions, media and professional authorities such as health professionals, is significantly lower than the European Union average. Such a combination of institutional distrust and susceptibility to disinformation, can call into question democratic standards, values and procedures.

Due to all this, Croatia must strengthen the resilience of society to disinformation, respecting the recommendations of all relevant institutions, initiatives and projects. Key activities in this process should include: 1) fostering quality and independent journalism, 2) improving professional standards in journalism, 3) establishing independent, specialised fact-checking projects, 4) strengthening trust in fact-checkers, 5) systematic empowering of citizens through development

of media literacy skills, 6) developing computer-aided solutions to combat disinformation, 7) strengthening cooperation of key stakeholders and strengthening synergy of all research, developmental, educational and fact-checking activities, for which it is essential to provide support from public funds, through a system of direct and indirect subsidies. This support should be based on clear and measurable project selection criteria in relation to the set objectives and related activities, while ensuring the transparency of the selection process at all its stages, transparent reporting on the beneficiaries and amounts allocated and monitoring the execution of the set goals and agreed activities.

Key international organisations such as the Council of Europe, the European Commission and UNESCO, as well as relevant projects such as the European Digital Media Observatory (EDMO), clearly state several key principles of establishing and strengthening fact-checking and information verification systems: project cooperation between different stakeholders, multidisciplinary teams and cooperation with the academic community; strengthening cooperation, complementarity and exchange between traditional media and independent fact-checking organisations; development and application of new technologies; open-source intelligence (OSINT); systematic education and workshops.

Based on the above principles, the **general objective** of the first part of the C1.1.1. R6-I2 "Establishment of Media Fact Checking and Public Disclosure System" investment, within the National Recovery and Resilience Plan, implemented by the Ministry of Culture and Media and the Agency for Electronic Media, should be to **strengthen the resilience of society to disinformation** by reducing the amount of disinformation and fake news in public space, strengthening information reliability and security in the consumption of media content and use of social networks, strengthening the quality of journalism and credible reporting, and strengthening media literacy.

The specific objectives of the programme should be a) to strengthen the competences of existing fact-checkers and develop new independent fact-checkers for checking the accuracy of information published in public space, media and social networks, b) to develop technology programmes and platforms for fact-checking, c) to strengthen media resilience to disinformation and encourage quality and credible journalism; d) to encourage media literacy projects based on education on combating disinformation; and e) to create a collaborative repository of checked information that should also serve as a collaborative platform for the exchange of tools, methodologies and knowledge in the field of combating disinformation and fact-checking.

In line with the stated objectives and based on the analysis presented in this Research Study, the investment should encourage the following activities (1) **strengthening the capacities and competencies of existing fact-checkers and the establishment of new independent fact-checkers**, namely fostering partnerships between media, civil society organisations and academic, educational and research institutions and fostering cooperation between media and fact-checking organisations; (2) **establishing and strengthening fact-checking systems and procedures in media newsrooms**, in particular encouraging the development of specialised fact-checking units within existing media newsrooms, especially through education and workshops for the use of new tools and technologies for fact-checking online materials and through the application of such tools in the work practices of newsrooms and through cooperation with independent fact-checking organisations/projects, in accordance with the recommendations of relevant international institutions; (3) **the development of computational fact-checking systems** and training and workshops on their application; (4) **the development of educational programmes and materials relat ed to fact-checking and combating disinformation and the development of media literacy**,

as the skill of critical understanding and responsible use of media and online platforms, among children, young people and citizens of all generations, in accordance with the *Conclusions of the EU Council on media literacy in a constantly changing world* (2020/C 193/06).

In order to ensure a synergistic effect and maximum social impact of the projects financed within this investment, it is recommended to establish a collaborative platform, modelled on similar examples of good practice (e.g., EDMO Truly Media). The platform should provide an open repository of fact-checks that will serve as a reference source for all media dedicated to credible and accurate reporting, researchers and other interested parties engaged in studying and combating disinformation. The platform should foster cooperation with international organisations and institutions dealing with this problem and provide space for the exchange and development of tools, methods and knowledge in the field of combating disinformation. Likewise, in the context of this investment, and following the European examples of good practice described in Chapter 3.2 (e.g., BBC and Channel 4 in the United Kingdom, Yle in Finland, NSK in Norway), the inclusion and contribution of public media to the process of empowering society to combat disinformation should be encouraged and envisaged.

In order to ensure the integrity of actions under the programme for establishing and strengthening the fact-checking system, it is essential that in the process of evaluating project proposals, a special criterion should evaluate the willingness and ability of applicants to comply with certain **principles of work** that should include independence, transparency, political and other impartiality, a commitment to open and fair correction of error and the right to publish corrections, the absence of conflicts of interest, strict avoidance of any form of discrimination and strict avoidance of the spread of hate speech.

Literature and sources

AEM (Agency for Electronic Media) (2013). Analiza Tv tržišta (*Eng. TV Market Analysis*) .AEM, Ipsos Puls, Media HUB, AGB Nielsen.

AEM (Agency for Electronic Media) (2016). Navike gledanja televizijskog programa. (*Eng. Habits of watching a television programme*). AEM and Ipsos Puls.

Agić, Ž. and Merkler, D. (2013, September). Three syntactic formalisms for data-driven dependency parsing of Croatian. *International Conference on Text, Speech and Dialogue* (pp. 560-567). Berlin, Heidelberg: Springer.

Agić, Ž., Ljubešić, N., Merkler, D. (2013, August). Lemmatization and morphosyntactic tagging of Croatian and Serbian. *Proceedings of the 4th Biennial International Workshop on Balto-Slavic Natural Language Processing* (pp. 48- 57).

Allen, J., Howland, B., Mobius, M., Rothschild, D., Watts, D. J. (2020). Evaluating the fake news problem at the scale of the information ecosystem. *Science Advances*, 6(14), eaay 3539.

Almer, I., Goluza, S., Mustapic, T. Sentence-Level Detection of Propaganda in News Articles. *Text Analysis and Retrieval 2020 Course Project Reports*, 1.

Amazeen, M. A. (2020). Journalistic interventions: The structural factors affecting the global emergence of fact-checking. *Journalism*, *21*(1), 95-111.

Anic, I., Bilic, I., Škudar, S. Bot Detection From a Single Tweet. *Text Analysis and Retrieval 2020 Course Project Reports*, 5.

Auer, S., Bizer, C., Kobilarov, G., Lehmann, J., Cyganiak, R., Ives, Z. (2007). Dbpedia: A nucleus for a web of open data. In *The semantic web*. Lecture Notes in Computer Science, vol 4825. (pp. 722-735). Springer.

Babić, K., Petrović, M., Beliga, S., Martinčić-Ipšić, S., Matešić, M., Meštrović, A. (2021). Characterisation of COVID-19-related tweets in the English language: a framework based on the Cro-CoV-cseBERT model. *Applied Sciences*, *11*(21), 10442.

Babić, K., Petrović, M., Beliga, S., Martinčić-Ipšić, S., Pranjić, M., Meštrović, A. (2021, May). Prediction of COVID-19 related information spreading on Twitter. In *2021 44th International Convention on Information, Communication and Electronic Technology* (MIPRO) (pp. 395-399). IEEE.

Bagić, D. and Šuljok, A. (2021). "Put a Mask and Keep a Distance" – A Study of Determinants of Protective Behavior against COVID-19 Disease in the Republic of Croatia *Sociologija i prostor: časopis za istraživanje prostornoga i sociokulturnog razvoja, 59*(219), 119- 147.

Barišic, L., Pisacic, M., Radovic, F. (2020). To Context or Not to Context? Analysis of De-contextualized Word Embeddings and Propaganda Detection Task. https://www.researchgate.net/profile/Filip-Radovic-2/publication/344492469_To_Context_or_Not_to_Context_Analysis_of_De-contextualized_Word_Embeddings_in_Propaganda_Detection_Task/links/5f7c4fbe92851c14bcb1788b/To-Context-or-Not-to-Context-Analysis-of-De-contextualized-Word-Embeddings-in-Propaganda-Detection-Task.pdf Accessed April 25, 2022.

Barrett, B. and Kreiss, D. (2019). Platform transience: changes in Facebook's policies, procedures, and affordances in global electoral politics. *Internet Policy Review*, *8*(4), 1-22.

Beck, B., Kanižaj, I. and Lechpammer, S. (2021). ClickBait: Ten Ways to Manipulate the Context. In J. Jurišić and Z. Hrnjić Kuduzović, (ed). *Media Agenda 2020-2030* (pp. 89-102). Hans Seidel Stiftung.

Bennett, W. L., and Livingston, S. (2018). The disinformation order: Disruptive communication and the decline of democratic institutions. *European Journal of Communication*, *33*(2), 122–139.

Bilić, P. (2012). Croatian Media System between Reflexive Modernization and Mediatization: Societal and Institutional Analysis. *Društvena istraživanja, 21 (4), 821-842.*

Bilić, P., Valečić, M., Prug, T. (2021). Monitoring media pluralism in the digital era. Application of the Media Pluralism Monitor in the European Union, Albania, Montenegro, the Republic of North Macedonia, Serbia & Turkey in the year 2020. European University Institute. https://cadmus.eui.eu/bitstream/handle/1814/71940/croatia_results_mpm_2021_cmpf.pdf?sequence=1&isAllowed=y Accessed April 25, 2022.

Bilić, P., Valečić, M., Prug, T. (2022). Monitoring media pluralism in the digital era. Application of the Media Pluralism Monitor in the European Union, Albania, Montenegro, the Republic of North Macedonia, Serbia & Turkey in the year 2021. European University Institute.

Biloš, A. (2020). Challenges of internet information space: how do Croatian citizens perceive fake news issue? *CroDiM*, 3(1), 166-185.

Birdsell, D. S. and Groarke, L. (1996). Toward a theory of visual argument. *Argumentation and Advocacy*, 33(1), 1–10.

Blom, J. N. and Hansen, K. R. (2015). Click bait: Forward-reference as lure in online news headlines. *Journal of Pragmatics*, 76, 87- 100.

Blumler, J. and Michael G. (1995). *The Crisis of Public Communication*, London: Routledge.

Bogović, P. K., Meštrović, A., Beliga, S., Martinčić-Ipšić, S. (2021). Topic modelling of Croatian news during COVID-19 pandemic. *2021 44th International Convention on Information, Communication and Electronic Technology* (MIPRO) (pp. 1044-1051). IEEE.

Bojanowski, P., Grave, E., Joulin, A., Mikolov, T. (2017). Enriching word vectors with subword information. *Transactions of the association for computational linguistics*, 5, 135-146.

Boltužić, F., and Schneider, J. (2016, Aug.). Fill the gap! analyzing implicit premises between claims from online debates. *Proceedings of the Third Workshop on Argument Mining* (ArgMining2016) (pp. 124-133).

Bovan, K., Baketa, N., Kovačić, M., Čorkalo Biruški, D. (2022). Trust us, we know what we are doing: institutional trust in Croatia during the COVID-19 crisis. *Southeast European and Black Sea Studies*.

Bradshaw, S. and Howard, P. N. (2019). The Global Disinformation Order: 2019 Global Inventory of Organised Social Media Manipulation. Working Paper 3. Oxford, UK: Project on Computational Propaganda.

Brautović, M. (2010). Development of Croatian Online Journalism 1993-2010 *Medianali*, 4 (8), 23-42.

Brogi, E., Carlini, R., Nenadić, I., Parcu, P. L., Viola de Azevedo Cunha, M. (2020). Monitoring media pluralism in the digital era: Application of the media pluralism monitor in the European Union, Albania and Turkey in the years 2018 2019: report 2020. European University Institute.

Bromley, M. (1997) 'The End of Journalism? Changes in Workplace Practices in the Press and Broadcasting in the 1990s'. In M. Bromley and T. O'Malley, (ed). *A Journalism Reader* (pp. 330-350). New York: Routledge.

Brüggemann, M. and Engesser, S. (2017). Beyond false balance: How interpretive journalism shapes media coverage of climate change. *Global Environmental Change*, 42, 58-67.

Casero-Ripolles, A. and Izquierdo-Castillo, J. (2013). Between Decline and a New Online Business Model: The Case of the Spanish News Industry. *Journal of Media Business Studies*, 10(1), 63–78.

Carlson, M. (2014). When news sites go native: Redefining the Advertising–editorial Divide in Response to Native Advertising. *Journalism*, 16 (7), 849 - 865.

Chadwick, A. (2017). *The Hybrid Media System: Politics and Power*. 2nd edition. Oxford: Oxford University Press.

Chan, M.S., Jones, C.R., Hall Jamieson, K. Albarracín, D. (2017). Debunking: A Meta-Analysis of the Psychological Efficacy of Messages Countering Misinformation. *Psychological Science*, 28 (11), 1531- 1546.

Chiou, L., and Tucker, C. E. (2018). Fake News and Advertising on Social Media: A Study of the Anti-Vaccination Movement. Working Paper No. 25223. Cambridge, MA: The National Bureau of Economic Research.

Cohen, N. S., Hunter, A., O'Donnell, P. (2019). Bearing the burden of corporate restructuring: Job loss and precarious employment in Canadian journalism. *Journalism Practice*, 13(7), 817-833.

Compton, J. R. and Benedetti, P. (2010). Labour, new media and the institutional restructuring of journalism. Journalism studies, 11 (4), 487- 499.

Council of the EU (2020). Council conclusions on media literacy in an ever-changing world 2020/C 193/06, ST/8274/2020/INIT, OJ C 193, 9.6.2020, p. 23–28

Da San Martino, G., Barrón-Cedeno, A., Wachsmuth, H., Petrov, R., Nakov, P. (2020, December). Semeval-2020 task 11: Detection of propaganda techniques in news articles. *Proceedings of the fourteenth workshop on semantic evaluation* (pp. 1377-1414).

Davis, M. (2010). "Why Journalism Is a Profession". In C. Meyers, (ed). *Journalism Ethics* (pp. 91-102). Oxford University Press.

Derczynski, L., Bontcheva, K., Liakata, M., Procter, R., Hoi, G. W. S., & Zubiaga, A. (2017, August).

Semeval-2017 Task 8: RumouReval: Determining rumour veracity and support for rumours. Proceedings of the 11th International Workshop on Semantic Evaluation (Semeval-2017) (pp. 69-76).

Deuze, M., Elefante, P., Steward, B. (2010). Media work and the recession. *Popular Communication*, 8(3), 226-231.

Devlin, J., Chang, M.W., Lee, K. & Toutanova, K. (2019). Bert: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of the 2019 NAACL, 4171–4186.

Di Buono, M. P. and Snyder, J. (2017). Linguistic Features and Newsworthiness: an Analysis of News style. *CLiC-it*.

Di Buono, M. P., Šnajder, J., Bašić, B. D., Glavaš, G., Tutek, M., Milić-Frayling, N. (2017, September). Predicting news values from headline text and emotions. *Proceedings of the 2017 EMNLP Workshop: Natural Language Processing meets Journalism* (pp. 1-6).

Di Buono, M. P., Tutek, M., Šnajder, J., Glavaš, G., Bašić, B. D., & Milić-Frayling, N. (2017, April). Two layers of annotation for representing event mentions in news stories. *Proceedings of the 11th Linguistic Annotation Workshop* (pp. 82-90).

Dimitrov, D., Ali, B. B., Shaar, S., Alam, F., Silvestri, F., Firooz, H., ... Da San Martino, G. (2021, August). SemEval-2021 Task 6: Detection of Persuasion Techniques in Texts and Images. *Proceedings of the 15th International Workshop on Semantic Evaluation* (Semeval-2021) (pp. 70-98).

Dobek-Ostrowska, B. (2015). 25 years after communism: Four models of media and politics in Central and Eastern Europe. In B. Dobek-Ostrowska, B. and M. Glowacki (ed). *Democracy and Media in Central and Eastern Europe 25 Years on* (pp. 11-45). Frankfurt: Peter Lang, pp. 11–45.

Dobek-Ostrowska B. (2019). How the Media Systems Work in Central and Eastern Europe. In E. Połońska and C. Beckett (ed.), Public Service Broadcasting and Media Systems in Troubled European Democracies (pp. 259-278). Cham: Palgrave Macmillan.

Dohle, S., Wingen, T., Schreiber, M. (2020). Acceptance and adoption of protective measures during the COVID-19 pandemic: The role of trust in politics and trust in science. *Social Psychological Bulletin*, 15(4), 1-23.

Elias, C. and Catalan-Matamoros, D. (2020). Coronavirus in Spain: Fear of 'Official' Fake News Boosts WhatsApp and Alternative Sources. *Media and Communication*. 8(2), 462–466.

EU Code of Practice on Disinformation (2018).
https://www.europeansources.info/record/eu-code-of-practice-on-disinformation/ Accessed April 25, 2022.

Eurobarometer (2019). Standard Eurobarometer 92.
https://europa.eu/eurobarometer/surveys/detail/2255. Accessed April 25, 2022.

Eurobarometer, Special 516 (2021) European citizens' knowledge and attitudes towards science and technology.
https://europa.eu/eurobarometer/surveys/detail/2237. Accessed April 25, 2022.

Eurobarometer (2021). Standard Eurobarometer 94. Kantar. https://europa.eu/eurobarometer/surveys/detail/2355 Accessed April 25, 2022.

Eurobarometer (2022). Standard Eurobarometer 96.
https://europa.eu/eurobarometer/surveys/detail/2553. Accessed July 15, 2022.

European Commission (2018a). Tackling online disinformation: a European approach Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. COM/2018/236. eur-lex.europa.eu/legal-content/en/txt/html/URI=CELEX:52018DC0236&from=en Accessed April 25, 2022.

European Commission (2018b). Action Plan against Disinformation. Joint Communication to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions. eur-lex.europa.eu/legal-content/en/TXT/HTML/?uri=CELEX:52018JC0036&from=en Accessed April 25, 2022.

European Commission (2020). Tackling COVID-19 Disinformation — Fact recognition — Getting the facts right. Joint Communication to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions. JOIN(2018) 36 final. JOIN/2020/8 final.

EUR-lex.europa.eu/legal-content/en/txt/html/? uri=CELEX:520JC0008&from=en Accessed April 25, 2022.

European Commission (2021). Guidance on Strengthening the Code of Practice on Disinformation. COM (2021) 262 final. https://digital-strategy.ec.europa.eu/en/library/guidance-strengthening-code-practice-disinformation

Faesen, L., Klimburg, A., van Hoeve, S., Sweijs, T. (2021). Red Lines & Baselines: Towards a European Multistakeholder Approach to Counter Disinformation. The Hague Center for Strategic Studies.
https://hcss.nl/report/red-lines-baselines/ Accessed April 25, 2022.

Gauch, S., Speretta, M., Chandramouli, A., Micarelli, A. (2007). User Profiles for Personalized Information Access. In P. Brusilovsky, A. Kobsa, W. Nejdl (ed). *The Adaptive Web: Methods and Strategies of Web Personalization* (pp. 54-89). Berlin: Springer.

Gillespie, T. (2018). *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*. New Haven: Yale University Press.

Glavaš, G. and Schneider, J. (2013, August). Recognizing identical events with graph kernels. *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics* (Volume 2: Short Papers) (pp. 797-803).

Glavaš, G. and Schneider, J. (2014). Event graphs for information retrieval and multi-document summarization. *Expert systems with applications*, 41(15), 6904-6916.

Gorrell, G., Kochkina, E., Liakata, M., Aker, A., Zubiaga, A., Bontcheva, K., Derczynski, L. (2019, June). SemEval-2019 task 7: RumourEval, determining rumour veracity and support for rumours. *Proceedings of the 13th International Workshop on Semantic Evaluation* (pp. 845-854).

Goyanes, M. (2014). An Empirical Study of Factors that Influence the Willingness to Pay for Online News. *Journalism Practice*, 8(6), 742–757.

Grave, É., Bojanowski, P., Gupta, P., Joulin, A., Mikolov, T. (2018, May). Learning Word Vectors for 157 Languages. *Proceedings of the Eleventh International Conference on Language Resources and Evaluation* (LREC 2018).

Graves, L. (2018). Boundaries not drawn: Mapping the institutional roots of the global fact-checking movement. *Journalism Studies*, 19 (5), 613- 631.

Graves, L. and Cherubini, F. (2016). The Rise of Fact-Checking Sites in Europe. Oxford: Reuters Institute for the Study of Journalism. https://reutersinstitute.politics.ox.ac.uk/our-research/rise-fact-checking-sites-europe Accessed April 25, 2022.

Grbeša, M. and Šalaj, B. (2018). *Dobar, loš ili zao?Populizam u Hrvatskoj. (Eng. The good, the bad or the evil? Populism in Croatia.*) Zagreb: Tim Press.

Grmuša, T. Prelog, L. (2020). The role of New Technologies in combatting Fake News — Experiences and Challenges of Croatian Media Organisations. *Media Studies*, 11 (22), 62-80.

Guo, Z., Schlichtkrull, M., & Vlachos, A. (2022). A survey on automated fact-checking. *Transactions of the Association for Computational Linguistics*, 10, 178-206.

Gupta, A., and Srikumar, V. (2021, August). X-Fact: A New Benchmark Dataset for Multilingual Fact Checking. *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing* (Volume 2: Short Papers) (pp. 675-682).

Hallin, D. C., and Mancini, P. (2004). *Comparing media systems: Three models of media and politics*. Cambridge: Cambridge University Press.

Hallin, D. C., and Mancini, P. (2012). *Comparing media systems beyond the western world*. Cambridge: Cambridge University Press.

Han, Q., Zheng, B., Cristea, M., Agostini, M., Bélanger, J. J., Gützkow, B., ... PsyCorona Collaboration. (2021). Trust in government regarding COVID-19 and its associations with preventive health behaviour and prosocial behaviour during the pandemic: a cross-sectional and longitudinal study. *Psychological medicine*, 1-11.

Helberger, N. (2016). Facebook is a new breed of editor: a social editor. Media Policy Blog, September 15, 2016 blogs.lse.ac.uk/medialse/2016/09/15/facebook-is-a-new-breed-of-editor-a-social-editor/ Accessed April 25, 2022.

Helberger, N. (2019). On the Democratic Role of News Recommenders. *Digital Journalism*, (7) 8, 993-1012. Henjak, A. (2017). Institutional Trust and Democracy Satisfaction in Croatia: Partisanship- Versus Outcome-Driven Evaluations. *Croatian and Comparative Public Administration*, 17 (3), 343-363.

Hermida, A. (2012). Tweets and Truth: Journalism as a discipline of collaborative verification. *Journalism Practice*, 6(5-6), 659-668.

Herrero, L., Humprecht, E., Engesser, S., Brüggemann, M., Büchel, F. (2017). Rethinking Hallin and Mancini Beyond the West: An Analysis of Media Systems in Central and Eastern Europe. *International Journal of Communication*, 11(27), 4797-4823.

Hill, S., and Lashmar, P. (2013). *Online journalism: The essential guide*. Sage.

HLEG (High Level Expert Group on Fake News and Online Disinformation). (2018). Final report. https://www.ecsite.eu/sites/default/files/amulti-dimensionalapproachtodisinformation-reportoftheindependenthighlevelgrouponfakenewsandonlinedisinformation.pdf Accessed April 25, 2022.

Holton, A. E. and Chi, H. AND. (2012). News and the Overloaded Consumer: Factors Influencing Information Overload Among News Consumers. *Cyberpsychology, behavior and social networking*, 15 (11), 619-24.

Horowitz, M., Cushion, S., Dragomir, M., Gutiérrez Manjón, S., Pantti, M. (2021). A Framework for Assessing the Role of Public Service Media Organizations in Counting Disinformation. *Digital Journalism*, 1-23.

HURA Media (2020). Praćenja medijskog pluralizma. (*Eng. Monitoring of media pluralism*). https://hura.hr/istrazivanja/medijska-potrosnja-u-hr/ Accessed April 25, 2022.

Immundo, M. N., and Rapp, D. N. (2021). When Fairness is Flawed: Effects of False Balance Reporting and Weight-of- Evidence statements on Beliefs and Perceptions of Climate Change. *Journal of Applied Research in Memory and Cognition*.

Ivanuš, Ž. (2021). Journalistic Ethics under Pressure from Interest Groups — Croatian Journalists' Experience. *Media, Culture and Public Relations*, 12(1), 73-90.

Karan, M., Glavaš, G., Šarić, F., Šnajder, J., Mijić, J., Šilić, A., Bašić, B. D. (2013). CroNER: Recognizing named entities in Croatian using conditional random fields. *Informatica*, 37(2).

Kiesel, J., Mestre, M., Shukla, R., Vincent, E., Adineh, P., Corney, D., ... Potthast, M. (2019, June). Semeval-2019 task 4: Hyperpartisan news detection. *Proceedings of the 13th International Workshop on Semantic Evaluation* (pp. 829-839).

Kietzmann, J., Lee, L. W., McCarthy, I. P., Kietzmann, T. C. (2020). Deepfakes: Trick or treat?. *Business Horizons*, 63(2), 135–146.

Kilicoglu, H., Shin, D., Fiszman, M., Rosemblat, G., Rindflesch, T. C. (2012). SemMedDB: a PubMed-scale repository of biomedical semantic predications. *Bioinformatics*, 28(23), 3158-3160.

King, G., Pan, J., Roberts, M. (2017). How the Chinese Government Fabricates Social Media Posts for Strategic Distraction, Not Engaged Argument. *American Political Science Review*, 111(3), 484–501.

Korenčić, D., Ristov, S., Šnajder, J. (2015, October). Getting the agenda right: measuring media agenda using topic models. *Proceedings of the 2015 Workshop on Topic Models: Post-Processing and Applications* (pp. 61-66).

Kovach, B. and Rosenstiel, T. (1999). *Warp Speed: America in the Age of Mixed Media*. New York: The Century Foundation.

Kovach, Bill and Rosenstiel, Tom (2001). *The Elements of Journalism*. New York: Crown Publishers.

Kung, L. (2015). Innovators in Digital News. (Extract). Reuters Institute for the Study of Journalism. reutersinstitute.politics.ox.ac.uk/sites/default/files/research/files/Innovators%2520in%2520Digital%2520News%2520Extract.pdf Accessed 25.04.2022.

Labaš, D. and Ciboci, L. (2011). The Index Affair — in the Focus of the Media and the Public. *MediAnali*, 5 (9), 1-40.

Laursen, J. and Stone, M. (2016). Native Advertising Trends 2016: The News Media Industry. Native Advertising Institute (NAI) and the International News Media Association (INMA). nativeadvertisinginstitute.com/wp-content/uploads/2016/10/TrendReportNewsMedia16.pdf Accessed 04/25/2022.

Lazarski, E. and Howard, C. (2021). Using NLP for Fact Checking: A Survey. *Designs*, 5(3), 42.

Lewis, S. C. (2012). The Tension Between Professional Control and Open Participation. Information. *Communication & Society*, 15(6), 836-866.

Lischka, J. A. and Garz, M. (2021). Clickbait news and algorithmic curation: A game theory framework of the relation between journalism, users, and platforms. *New Media & Society*, 1461444821102717 4.

Lundby, K. (2014). *Mediatization of Communication: Handbooks of Communication Science*. Berlin: De Gruyter Mouton.

Luttenberger, L. and Vulinovic, K. Claim Strength Identification for Detecting Exaggerations in Science News. *Text Analysis and Retrieval 2018 Course Project Reports*, 75.

Lyons, B., Mérola, V., Reifler, J., Stoeckel, F. (2020). How politics shape views toward fact-checking: Evidence from six European countries. *The International Journal of Press/Politics*, 25(3), 469-492.

Ljubešić, N. and Dobrovoljc, K. (2019, August). What does neural bring? Analysing improvements in morphosyntactic annotation and lemmatisation of English, English and English. *Proceedings of the 7th workshop on balto-slavic natural language processing* (str. 29-34).

Ljubešić, N. and Erjavec, T. (2011, September). hrWaC and slWaC: Compiling web corpora for Croatian and Slovene. *International Conference on Text, Speech and Dialogue* (pp. 395-402). Berlin, Heidelberg: Springer.

Ljubešić, N. and Lauc, D. (2021, April). BERTić-The Transformer Language Model for Bosnian, Croatian, Montenegrin and Serbian. *Proceedings of the 8th Workshop on Balto-Slavic Natural Language Processing* (pp. 37-42).

Ljubešić, N., Stupar, M., Jurić, T., Agić, Ž. (2013). Combining available datasets for building named entity recognition models of Croatian and Slovene. *Slovenščina*, 2(1), 35-57.

Matteo, S. and Zotto, C. D. (2015). Native Advertising, or How to Stretch Editorial to Sponsored Content Within a Transmedia Branding Era. In G. Siegert, K. Forster, S. Chan-Olmsted, M. Ots, editor (s). *Handbook of Media Branding* (pp. 169-185). Cham: Springer.

Mihaylova, T., Karadzhov, G., Atanasova, P., Baly, R., Mohtarami, M., Nakov, P. (2019, June). Semeval-2019 Task 8: Fact Checking in Community Question Answering Forums. *Proceedings of the 13th International Workshop on Semantic Evaluation* (pp. 860-869).

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems*, 26.

Mozur, P., Kao, J., Krolik, A. (2020). No 'Negative' News: How China Censored the Coronavirus. *The New York Times* (December 19). https://www.nytimes.com/2020/12/19/technology/china-coronavirus-censorship.html Accessed April 25, 2022.

Nenadić, I. (2019). Unpacking the "European approach" to tackling challenges of disinformation and political manipulation. *Internet Policy Review*, (8)4, 1-22.

Nenadić, I. (2020a). Twitter and changing journalistic practice in Croatia. Doctoral dissertation. Josip Juraj Strossmayer University of Osijek.

Nenadić, I. (2020b). Power without responsibility: content moderation policies of online-platforms in the response to infodemic. *Anali hrvatskog politološkog društva*, 17, 79-102.

Nenadić, I. (2021). EC's Guidance to Strengthening the Code of Practice on Disinformation: A Mis-take with Mis-information?. Discussion Series, Centre for Media Pluralism and Media Freedom, European University Institute. https://cmpf.eui.eu/ecs-guidance-to-strengthening-the-code-of-practice-on-disinformation-a-mis-take-with-mis-information/ Accessed April 25, 2022.

Nenadić, I. and Ostling, A. (2018). Media innovation in Europe and reinvention of audiences: between citizens and consumers. *Media Studies*, 9 (17), 4-22.

Nenadić, I. and Vučković, M. (2021). Disinformation: Educational brochure and exercises for understanding disinformation problems. Agency for Electronic Media and UNICEF.

Nieminen, S. and Rapeli, L. (2019). Fighting misperceptions and doubting journalists' objectivity: A review of fact-checking literature. Political Studies Review, 17(3), 296-309.

Ó Fathaigh, R., Helberger, N., Appelman, N. (2021). The perils of legally defining disinformation. *Internet Policy Review*, 10(4).

Oulasvirta, A. and Blom, J. (2007). Motivations in Personalisation Behaviour. *Interacting with Computers*, 20(1), 1–16.

Palić, N., Vladika, J., Čubelić, D., Lovrenčić, I., Buljan, M., Šnajder, J. (2019, June). TakeLab at SemEval-2019 Task 4: Hyperpartisan News Detection. *Proceedings of the 13th International Workshop on Semantic Evaluation* (pp. 95-998).

Pariser, E. (2011). *The Filter Bubble: What the Internet is Hiding from You.* UK: Penguin.

Peruško, Z. (2012). Media system in Croatia: from Authoritarian to the Mediterranean model. U V.

Peruško, Z. (2013). Rediscovering the Mediterranean characteristics of the Croatian media system. *East European Politics and Societies*, 27(4), 709-726.

Peruško, Z. (2016). Historical institutionalist approach in comparative media systems research: The case of post-Yugoslavia. Javnost – The Public, 23(3), 255–272.

Peruško, Z. (2021). Public Sphere in Hybrid Media Systems in Central and Eastern Europe. *Javnost - The Public*, 28(1), 36-52.

Peruško, Z. and Čuvalo, A. (2014). Comparing socialist and post-socialist television culture: Fifty years of television in Croatia. *View – Journal of European Television & Culture*, 3 (5), 131–150.

Peruško, Z., Perišin, T., Topić, M., Vilović, G., Zgrabljić Rotar, N. (2011). *Croatian media system according to UNESCO indicators of media development*. Zagreb: Croatian Political Science Library, Faculty of Political Science.

Peruško, Z., Čuvalo, A., Vozab, D. (2016). Country Report. Journalists in Croatia. The Worlds of Journalism Study. https://epub.ub.uni-muenchen.de/29703/1/Country_report_Croatia.pdf Accessed April 25, 2022

Peruško, Z., Vozab, D., Čuvalo, A. (2021). *Comparing Post-Socialist Media Systems: The Case of Southeast Europe*. Abingdon: Routledge.

Peters, M., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., Zettlemoyer, L. (2018). Deep contextualized word representations. *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Volume 1 (Long Papers) 2227–2237.

Picard, R. G. (2014). New Approaches to Paid Digital Content. Digital News Report 2017. Reuters Institute for the Study of Journalism, Oxford University. https://www.digitalnewsreport.org/essays/2014/new-approaches-to-paid-digital-content/ Accessed April 25, 2022.

Poliak, A. (2020). A survey on recognizing textual entailment as an NLP evaluation. arXiv preprint arXiv:2010.03061.

Puljiz, S. Ravlić, V. Visković (ed). *Hrvatska U Europskoj uniji:kako dalje? (Eng. Croatia in the European Union: how to proceed?)* (pp. 437-475). Zagreb: Centar za demokraciju i pravo Miko Tripalo.

Pušelj, D. and Škalec, T. Propaganda in Press: Challenges of Automatic Detection. *Text Analysis and Retrieval 2020 Course Project Reports*, 55.

RIDNR (Reuters Institute Digital News Report). (2022). https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2022. Accessed June 25, 2022.

RIDNR (Reuters Institute Digital News Report). (2021). https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2021 Accessed April 25, 2022.

RIDNR (Reuters Institute Digital News Report). (2020). https://www.digitalnewsreport.org/survey/2020/ Accessed April 25, 2022.

RIDNR (Reuters Institute Digital News Report). (2019). https://www.digitalnewsreport.org/survey/2019/ Accessed April 25, 2022.

RIDNR (Reuters Institute Digital News Report). (2018). https://reutersinstitute.politics.ox.ac.uk/our-research/digital-news-report-2018 Accessed April 25, 2022.

RIDNR (Reuters Institute Digital News Report). (2017). https://reutersinstitute.politics.ox.ac.uk/our-research/digital-news-report-2017 Accessed April 25, 2022.

Saltz, E., Leibowicz, C. R., Wardle, C. (2020). Encounters with visual misinformation and labels across platforms: An interview and diary study to inform ecosystem approaches to misinformation interventions. Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems, pp. 1-6.

Samuels, G. (2017). The BBC has started a team to take down 'fake news'. *The Independent* (January 13). www.independent.co.uk/news/media/bbc-team-debunk-fake-news-stories-media-james-harding-a7525686.html Accessed April 25, 2022.

Singer, J. B. (2004). Strange bedfellows? The diffusion of convergence in four news organizations. *Journalism studies*, 5(1), 3-18.

Soares, F., and Recuero, R. (2021). How the Mainstream Media Help to Spread Disinformation about Covid-19. *M/C Journal*, 24(1).

Storks, S., Gao, Q., Chai, J. Y. (2019). Recent Advances in Natural Language Inference: A Survey of Benchmarks, Resources, and Approaches. arXiv preprint arXiv:1904.01172.

Strömback, J. and Frank, E. (2009). Shaping Politics: Mediatization and Media Interventionism. In K. Lundby, (ed.). *Mediatization. Concept, Changes, Consequences* (pp. 205–224). NY: Peter Lang.

Sunstein, C. R. (2002). *Republic.com*. Princeton: Princeton University Press.

Sunstein, C. R. (2009). *Going to Extremes: How Like Minds Unite and Divide*. New York: Oxford University Press.

Swire-Thompson, B., DeGutis, J., Lazer, D. (2020). Searching for the Backfire Effect: Measurement and Design Considerations. *Journal of applied research in memory and cognition*, 9(3), 286–299.

Ružica Šimić Banović (2019) Uhljeb – a post-socialist homo croaticus: a personification of the economy of favours in Croatia?, Post-Communist Economies, 31:3, 279-300, DOI: 10.1080/14631377.2018.1537708

Šnajder, J., Bašić, B. D., Tadić, M. (2008). Automatic acquisition of inflectional lexica for morphological normalisation. *Information Processing & Management*, 44(5), 1720-1731.

Švob-Đokić, N., Bilić, P., Peruško, Z. (2011). Does media policy promote media freedom and independence? The case of Croatia. https://www.eliamep.gr/wp-content/uploads/2012/02/Croatia.pdf Accessed April 25, 2022.

Tandoc, Jr., E. C., Lim, Z. W. Ling, R. (2017). Defining 'Fake News': A Typology of Scholarly Definitions. *Digital Journalism*, 5(7), 1–17.

Tandoc Jr, E. C., Ling, R., Westlund, O., Duffy, A., Goh, D., Zheng Wei, L. (2018). Audiences' acts of authentication in the age of fake news: A conceptual framework. *New media & society*, 20(8), 2745-2763.

The World Press Freedom Index (2022). Data for Croatia. https://rsf.org/en/croatia Accessed April 25, 2022.

Thorne, J. and Vlachos, A. (2018). Automated fact checking: Task formulations, methods and future directions. arXiv preprint arXiv:1806.07687.

Thorne, J., Vlachos, A., Cocarascu, O., Christodoulopoulos, C., Mittal, A. (2018). The Fact Extraction and VERification (FEVER) Shared Task. *Training*, 80 (29,775), 35-639.

Tonković, M., Dumančić, F., Jelić, M., Čorkalo Biruški, D. (2021). Who believes in COVID-19 conspiracy theories in Croatia? Prevalence and predictors of conspiracy beliefs. *Frontiers in Psychology,* 12.

Tsfati, Y., Boomgaarden, H. G., Strömbäck, J., Vliegenthart, R., Damstra, A., Lindgren, E. (2020). Causes and consequences of mainstream media dissemination of fake news: literature review and synthesis. *Annals of the International Communication Association*, 44(2), 157-173.

Ulčar, M. and Robnik-Šikonja, M. (2020, May). High Quality ELMo Embeddings for Seven Less-Resourced Languages. *Proceedings of the 12th Language Resources and Evaluation Conference* (pp. 4731-4738).

Ulčar, M. and Robnik-Šikonja, M. (2020, September). Finest bert and crosloengual bert. *International Conference on Text, Speech, and Dialogue* (pp. 104-111). Cham: Springer.

Van Dijck, J. (2020). Governing digital societies: Private platforms, public values. *Computer Law & Security Review*, 36, 105377

Vesnic-Alujevic, L., Jurišić, J., Bonacci, D. (2021). Changing sourcing practices: Journalists' use of Facebook in Croatia. *Journal of Applied Journalism & Media Studies*, 10(3), 253-273.

Volarević, M. and Bebić, D. (2013). Social Networks as News Sources in Croatia's most viewed Television Newscasts. *Medijske studije*, 4 (8), 60-75.

Vosoughi, S., Roy, D., Aral, Si. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151.

Vrtič, I. and Car, V. (2016). Manipulation in media discourse: critical discourse analysis of the headlines on Croatian President Kolinda Grabar-Kitarović published in the most viewed Croatian news websites. In V. Car, M. Matović, L. Turčilo, (ed). *Mediji i javni interes* (pp. 145-165). Zagreb: Faculty of Political Science and Hanns Seidel Stiftung.

Walter, N. and Tukachinsky, R. (2020). A Meta-Analytic Examination of the Continued Influence of Misinformation in the Face of Correction: How Powerful Is It, Why Does It Happen, and How to Stop It?. *Communication Research*, 47(2), 155-177.

Wang, N. X., Mahajan, D., Danilevsky, M., Rosenthal, S. (2021, August). Semeval-2021 Task 9: Fact Verification and Evidence Finding for Tabular Data in Scientific Documents (SEM-TAB-FACTS). *Proceedings of the 15th International Workshop on Semantic Evaluation* (Semeval-2021) (pp. 317-326).

Wardle, C. (2018). 5 Lessons for Reporting in an Age of Disinformation. First Draft (December 27). https://firstdraftnews.org/articles/5-lessons-for-reporting-in-an-age-of-disinformation/ Accessed April 25, 2022.

Wardle, C. and Derakhshan, H. (2017). Information disorder: Toward an interdisciplinary framework for research and policymaking. Report No. DGI(2017)09. Strasbourg: Council of Europe. https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-researc/168076277c accessed April 25, 2022.

WHO (World Health Organization). (2020). Novel Coronavirus (2019- nCoV). Situation Report — 13 (February 2, 2020). https://apps.who.int/iris/handle/10665/330778 Accessed April 25, 2022.

Zeng, X., Abuansour, A. S., Zubiaga, A. (2021). Automated Fact-Checking: A Survey. arXiv e-prints, arXiv-2109.

Zhong, R., Mozur, P., Kao, J., Krolik, A. (2020). No 'Negative' News: How China Censored the Coronavirus. *The New York Times* (December 19). https://www.nytimes.com/2020/12/19/technology/china-coronavirus-censorship.html accessed April 25, 2022.

Zuiderveen Borgesius, F., Trilling, D., Möller, J., Bodó, B., De Vreese, C. H., Helberger, N. (2016). Should we worry about filter bubbles?. *Internet Policy Review*, 5(1).

Glossary

ALGORITHMS — a finite series of precisely defined computationally feasible steps. In essence, all the tasks performed by the computer are reduced to performing the instructions given in the form of an algorithm. In addition to the term "algorithm", literature in the Croatian language sometimes uses the term "procedure". In the context of trends in the modern communication environment, the term *algorithms* often refers to computational functions that are essential for the organisation of a large amount of information on the Internet. These computational functions, among other things, adapt the content offered to users based on their preferences and previous browsing of the Internet and social networks.

ATTENTION ECONOMY — in the context of information overload, attention is one of the key scarce economic resources today. For this reason, a whole range of economic and political activities are aimed at developing strategies, tools and techniques to win, retain and capitalise on human attention. The term is used in political communication and media studies to explain the practices that digital media (social networks, applications, etc.) continuously seek to "hijack" and retain the attention of users.

BOTS — robots, i.e., software applications that perform simple and repetitive automated network tasks by mimicking human behaviour but acting much faster than a human could. Social media bots try to look like normal profiles; they act mainly by commenting, liking and sharing certain content to enhance its reach and create a semblance of relevance.

CLICKBAIT — the practice of writing sensationalist and bombastic headlines in which information is deliberately withheld in order to "lure" audiences to open an article by clicking on the title.

DEBUNK — the process and result of revealing a false claim or content.

DISINFORMATION — verifiably false or misleading information designed, presented and distributed for economic, political or other gain with the intention of misleading the public.

FACT-CHECKING — a systematic approach to checking information. Unlike journalism that should check information before publication, fact-checkers check the information that has already been published and circulated in public space. *Fact-checking* organisations often act as independent organisations and projects. They are often associated with civil society organisations and bring together journalists with previous experience of working in the media. There are more and more examples of specialised fact-checking departments within media organisations to strengthen the process of verifying sources and, in general, checking information before it is published.

FAKE NEWS — "viral posts based on fictitious accounts made to look like news reports" (Tandoc et al., 2017) with the intention to deceive readers for financial, ideological or other gain. A number of authors and relevant organisations suggest how the term "fake news" is inadequate as it fails to effectively encompass the complexity of various dimensions of information disorder, where the contents are not always completely "false", but the fabricated information is sometimes mixed with facts, taken out of context, some key information is deliberately omitted, etc. Similarly, the contents are not always in the form of news but also include various forms of manipulated video materials, memes, infographics, microtargeting, organised trolling, bot networks, etc. In addition, the term "fake news" has been politicised and used to discredit the media and other institutions.

FALSE BALANCE — in the context of information disorder and relativisation of science and scientific evidence, the journalistic standard of balanced reporting, i.e., balanced presentation of different sides of the topic has also been affected. Misinterpretation of this standard has resulted in a *false balance*, that is, giving unfounded equal media space to opposing sides in a debate. (e.g., on climate change), although these sides are not founded on the same level of scientific basis or supported equally by scientific evidence (Immundo and Rapp, 2021).

INFODEMIC — occurs during high-intensity events when, in a short period and due to great public interest, the amount of information increases greatly, making it difficult to distinguish rumours and disinformation from credible news. The term was promoted by the World Health Organization in 2020 in the context of the COVID-19 pandemic.

INFORMATION DISORDER – evidence suggests that the business models of online platforms and other digital intermediaries, which have become the main source of news and information for many, facilitate or even encourage the dissemination of sensationalist, misleading and unreliable media content. In information disorder, disinformation, misinformation and malinformation are mixed. In order to fully understand the phenomenon, we should consider the actors and their motivation, the contents and the way they are formed, the factors that make certain individuals or groups more vulnerable to manipulation and the mechanisms of dissemination and visibility of such contents.

INFORMATION OVERLOAD — the emergence of the Internet, and especially the rise of online platforms, has enabled the dissemination of a large amount of information often managed by algorithms, lacking editorial oversight and transparency of content display and ranking. Information overload affects people's focus and attention, making it much harder for many to identify and access quality journalism.

MALINFORMATION — information that is based on reality, but used maliciously to inflict harm on a person, organisation or country. An example would be sharing someone's intimate photos and videos without their consent.

MISINFORMATION — information that is false but not created with the intention of causing harm. This is misleading or inaccurate information that people share, believing that it is accurate. Unintentional journalistic errors are also considered to be misinformation.

NATURAL LANGUAGE PROCESSING — a branch of artificial intelligence that combines machine learning methods with the knowledge of linguistics and related fields. It deals with the research of procedures for computational processing of natural language data that most often comes in the form of text.

SOCIAL NETWORKS — a common term generally referring to online platforms that enable users to connect, network, communicate, inform, entertain and share information. There are different users of social networks and their motivations vary. They include individuals of different generations who mostly share personal content and views, the media and journalists, as well as corporations and political actors who use social networks strategically paying for advertising to achieve certain goals. Some of the most famous social networks are Facebook, Twitter, LinkedIn. The term is often used for platforms that are focused on visual content, such as Instagram, TikTok, YouTube, etc.

ONLINE PLATFORMS — the development of the web infrastructure, content and applications has created a new area of moderation between natural persons or between natural and legal persons based on the processing of large amounts of (personal) data and personalisation of experience or offer. Online platforms cover a wide range of activities (e.g., Airbnb, Uber), but in the context of this Research Study the term refers to intermediaries that distribute, moderate and rank content, mainly through algorithmic processing, such as social networks (e.g., Facebook, Twitter), platforms designed primarily for the exchange of video content (e.g., TikTok, YouTube) and search engines (e.g., Google).

MACHINE LEARNING — a branch of artificial intelligence focused on designing algorithms which improve their efficiency based on empirical data.

ARTIFICIAL INTELLIGENCE — a part of computer science which deals with the development of the ability of the computer to perform tasks that require some form of intelligence, such as managing new opportunities, learning new concepts, making conclusions, understanding natural language or recognising scenes.

AGENCY FOR
ELECTRONIC
MEDIA