THE SOCIAL POTENTIAL OF FACT-CHECKING IN VERIFYING INFORMATION ABOUT THE WAR IN UKRAINE - A CASE STUDY

Joanna Urbaś[1]

Abstract

The emergence of new media has turned the formerly passive recipients of content into active users who, as a part of various initiatives, are willing to establish virtual communities. The aim of the argument is to present the social dimension of fact-checking in social media through the example of the analysis of the profile of the Demagog Association in the first month after the outbreak of the war. The author will attempt to answer questions about the social nature of fact-checking - what topics and values the community in the Demagog profile gathers around, whether fact-checking has social potential and whether the community around fact-checking organizations has the potential to educate and activate the rest of society in social media.

Key words: fact-checking, fake news, post-truth, Demagog, social media

Introduction

The transition from passive observation to active participation related to content creation is a characteristic element of new media, closely intertwined with digitization. This transformation has not only accelerated the dissemination and reception of information but has also enabled the exchange of content and influence over it (Szymkiewicz, 2022, pp. 45-46). One noticeable phenomenon in this context is the pursuit of fact-checking as means to verify the accuracy of media messages. This phenomenon aims to establish a discipline of content verification in the infosphere, primarily on the Internet. It involves the verification of facts, data, and the accuracy of claims circulating in the public sphere (Szymkiewicz, 2022, p. 49). What are the implications of this and where do they result from? What is the social potential of this phenomenon and what risks does it entail? This article attempts to address these questions. Theoretical concepts related to the topic will be juxtaposed with an analysis of the content of a selected fact-checking profile. The issues raised in this work can serve as a starting point for further exploration in future research.

The study of the changes occurring in a previously well-explored issue regarding the role of media, should be started by reflecting on the role of media in a democratic

1   PhD candidate, Pontifical University of John Paul II in Cracow, e-mail: joanna.urbas@doktorant.upjp2.edu.pl; ORCID: 0000-0001-9475-3326

system. In the publication *Mass Media: Art, Entertainment, Business,* Mrozowski highlights the role of media in a liberal democracy model. He writes that media's stimulation of communicative interactions among all social actors leads to the activation of society and triggers social-political engagement. (Mrozowski, 2001, pp. 122-124). Current research indicates that social media, too, can serve as a tool for communication between citizens and public entities (Allcott, Braghieri, Eichmeyer, Gentzkow, 2020).

While the researcher was writing about mass media, social media can fulfill a similar role. Communication through them is a significant feature, and acquainting oneself with information through fact-checking organization profiles can contribute to community building.

Thus, in order to consider the phenomenon of social fact-checking, a definition of social media needs to be provided. In 2010, Kaplan and Haenlin identified two key characteristics defining social media. According to them, social media are "internet applications based on the ideological and technological foundations of Web 2.0, which allow for the creation and exchange of user-generated content" (Kaplan, Haenlein, 2010, pp. 59-68). In more recent definitions, authors indicate that social media are a digital space created by people and for people, providing an environment conducive to interactions and connections on various levels, including personal, professional, business, political, and social dimensions (Kapoor et al., 2018).

Theorists emphasize three components that play an equally important role and are referred to as the social media triangle: social networks, content, and Web 2.0. Specific content is not only created by community members but also shared with a broader audience, modified, and commented on. This includes photos, videos, information, reviews, comments, and more (Popiołek, 2015, pp. 60-71). "The content of websites is generated and modified most often by users (user-generated content). Creating this content is made possible by the specific rules of Internet functioning, known as Web 2.0. This concept is mainly associated with content generated by users. But as Dominik Kaznowski points out, Web 2.0 has a slightly broader meaning, especially in the common discourse, and also refers to the ways technology is used" (Popiołek, 2015, pp. 60-71). The quoted definitions underscore the dimension of content creation and sharing by users. This multi-directional communication sets new media apart from traditional media. Key attributes which had been ascribed by researchers to new media include interconnectivity, users acting as both senders and receivers, interactivity, a multitude of use cases, openness, ubiquity, spatial indefiniteness, and delocalization (Skrzypczak, 1999, p. 376). All these characteristics enable the formation of communities around the phenomenon of fact-checking.

The network, however, continues to evolve, contributing to the distinction of successive generations of networks. In Web 3.0, artificial intelligence mechanisms have been enhanced, especially in the aspects of data segregation and transmission (Tasner et al., 2011). In the further evolution of the network, referred to as Web 4.0 or the symbiotic network, a period is anticipated for the so-called Internet of Things and the symbiosis between humans and computers (Aghaei, 2012). Web 5.0, on the other hand, aims to implement sensory solutions from the field of neurotechnology, capable of recording and analyzing users' emotions in real-time (Chomiak-Orsa, Smoląg, 2021).

Virtual communities are a collective of individuals whose interactions take place through the Internet, hence the use of the term "online community". (Bonek i Szabo, 2013) Online communities primarily revolve around similarities, such as interests, opinions, or values. Members participate in them intentionally, having full freedom and flexibility in their engagement. Furthermore, these communities are non-spatial and asynchronous, which promotes intentionality as interactions are not limited by space or time (Siuda, 2006, pp. 179-185). The author hypothesizes that the fact-checking community also gathers around similarities, primarily in interests and values. These similarities include a shared interest in contemporary world events and values such as truth, honesty,

and critical thinking.

Participation in new virtual groups allows individuals to decide on their affiliation with these groups. These are referred to as new nomads, characterized by the absence of traditional structures in terms of stability. They are based on emotions, tastes, interests, or common pursuits. (Mazur, Kuć, 2020, pp. 165-183) The community gathering around fact-checking profiles can thus be associated with one of the classical types of virtual communities – a community of interests, distinguished by shared hobbies and a higher degree of communication than, for example, a transactional community where people focus solely on buying and selling (Armstrong, Hagel, 1996).

Undoubtedly, new media, including social media, with their openness and interactivity, offer many opportunities for the development of civic engagement. One of such initiatives can be fact-checking activities. Despite the fact that, as indicated in the *Disinformation Through the Eyes of Poles* report, only 5% of respondents use fact-checking portals as a source of information, and 12% declare that they visit them when they want to verify information, attention is drawn to the increasing number of initiatives related to fact-checking. Examples of such activities that emerged after February 24th include the activities on social media by the state research institute NASK (National Academic Computer Network) as part of the *Engage in Verification* project and the initiative by the company Brand24, *Debunking Disinformation*.

An inherently related concept in the realm of fact-checking is fake news. The term "fake news" refers to media messages that are neither true nor false, often based on misinformation, and frequently containing elements of truth (Gillin, 2017). Yuval Harari expands the definition, describing fake news as a deliberate implementation of falsehood in the public domain, not limited to the media space (Harari, 2018). In many definitions (Elliot and Culver, 1992; Gans, 2004; Allcott and Gentzkow, 2017), the aspect of conscious creation of fake news to shape public attitudes is emphasized.

According to a study conducted in 2019 on a sample of 1,000 individuals aged 15 and above, the primary responsibility to combat disinformation on the Internet aimed at influencing the outcomes of democratic elections in Poland, should lie with website, portal, and app administrators, as believed by 55% of respondents. These are followed by the police, the prosecutor's office (27.5%) and the Polish government with its central offices (26.6%). The significant difference in percentage points between the first position and the subsequent ones signals that those recipients have high expectations regarding those responsible for managing online content and consider them to bear the greatest responsibility for presenting accurate information (Bochenek, Lange, 2019, p. 7).

The first organized forms of such activity began to emerge after 2000, with Fact-Check.org being an example, launched in 2003 to monitor political issues (Łódzki, 2017, pp. 19-28). In Poland, the first fact-checking organization is the Demagog Association, which has been a member of the International Fact-Checking Network (IFCN) since May 2019. As a member of this network, the Association is obliged to adhere to principles such as impartiality and fairness, source transparency, funding and organization transparency, methodological transparency, and a policy of open and honest corrections.

Under the "Methodology" section on the Association's website, there is a description of the fact-checking process. It involves the analysis of all statements that can be fact-checked without random selection – more frequently, objectively verifiable political statements appear on the website. Sources include recordings of radio and television programs, transcripts of parliamentary sessions, direct statements by politicians on social media, as well as requests from readers submitted through forms or social media profiles. The last element is particularly important in the context of this article because it demonstrates the community-oriented nature of fact-checking.

Research Methodology

The aim of the study is to determine the potential for the formation of a community around the phenomenon of fact-checking using the Demagog Association's Facebook profile as an example. The main research question posed by the article's author is: What are the ways in which the social character of fact-checking is implemented on Facebook? Specific research questions concern the categories of content that the Demagog profile's community gathers around and the roles played by various social behaviors. The research is exploratory and can contribute to the development of a tool for future research initiatives.

The author opted to analyze the profile of the Demagog Association due to its status as the first fact-checking organization in Poland. Additionally, it is a particularly active organization on social media platforms, having amassed a significant number of followers.

The research was conducted in the following stages: initially, the author isolated the activity in the profile for a specific period. The selected period was one month from February 24th to March 24th, 2022, as this was the first month of Russia's aggression against Ukraine, which was associated with a particularly high intensity of misinformation activities on the internet and increased vigilance among fact-checkers. Social media are the primary source of information for 54% of individuals aged 18-34 (Digital Poland, 2022). During times of war, this is particularly associated with an increased risk of misinformation (Institute of Central Europe 2022).

The author aims to draw attention to the aspect of the social media war, as the conflict described here is increasingly referred to as the first social media war. (Suciu, 2022) This is because the involved parties consistently publish frontline reports on social media platforms, followed by rapid sharing by users. Consequently, there is an enhanced potential for the dissemination of disinformation, underscoring the particularly crucial role of fact-checking organizations during this period.

Basic quantitative analysis was also conducted to present numerical data related to interactions in various thematic content categories. The author analyzed posts, rather than original publications on the Demagog Association's website, because she believed that it was these posts, through their presence on social media, that had a greater potential to build a community around fact-checking.

The author, within the framework of her research, employs the method of content analysis. The origins of the method can be tracked by some scholars in Medieval analyses of the Bible conducted by theologians. (Krippendorff, 2003) The content analysis is described by Walery Pisarek, "a set of various techniques for the systematic study of streams or sets of messages, consisting of the objective (in practice, usually intersubjectively consensual) identification and categorization of their as unambiguously formal or content-related elements and an approximate (usually quantitative) estimation of the distribution of these elements and mainly comparative inference, aiming at understanding the content of the messages and other elements and conditions of the communication process" (Pisarek, 1983, s. 43). The application of this method for the purposes of this article is justified by the fact that, as Lisowska-Magdziarz indicates, it is a method that is particularly suited for chaotic and unorganized material, which social media content undoubtedly represents (Lisowska-Magdziarz, 2004).

The criteria for the content analysis include qualitative criteria such as the category of social behavior, form of interaction, and type of comment, as well as quantitative criteria such as the number of posts on a given day, the number of reactions, and the number of comments. Subsequently, the author identified social behavior in the comments section using her own categorization and provided interpretation for it. The corpus of the analyzed comments comprises of 15,209 entries categorized and based on data from the study, capturing their structure and enabling the description of their role in commu-

nity building.

The author identifies a gap in the available research related to fact-checking. Therefore, the article is intended to provide a new perspective on the subject with an emphasis on the community. The knowledge derived from the research may have value not only for new media researchers and sociologists but also practical significance for administrators of profiles dealing with information verification.

The research finding

Social behaviors are associated with interactions between users or between users and profile administrators. Therefore, the author identifies the following indicators of the social nature of fact-checking: reactions to posts, comments, and shares. These are the basic forms of community communication available on the Facebook platform. While they do not exhaust the full potential of Facebook, these are the most observable forms, especially in the case of the Demagog profile.

The study identified two key categories of social fact-checking behaviors, which are evident in the comments on the Demagog profile: active and passive actions. Active actions include comments on posts and sharing posts, while passive actions involve observing the profile and reacting to posts. Active actions provide an opportunity to convey information from the post to other users (mainly friends) or reflect greater engagement through commenting. The mere observation of the profile and clicking reactions to posts, is contributing to fact-checking, but is still considered as passive due to the lower user engagement and lesser contribution to the viral spread of content.

From the analysis of comments, it can be concluded that they can be classified into the following thematic categories- as shown in Table 1. Below - along with descriptions, examples, and an attempt to outline their functions.

Type of Comment	Characteristics	Example (comment in the original spelling and translation)	Role for the Community
Approving Comment	Expression of support or approval for the post's content or Demagog's profile, thanking for the publication	„Demagog świetna robota, szerujemy link w anglojęzycznym internecie gdzie tylko się da!!" 27.02 "Demagog, great job, we'll share the link in the English-speaking Internet wherever we can!!" 27.02	Draws attention to content important for the user - when it appears among the most popular comments, it can act as social proof of correctness (Cialdini 2013).
Criticizing Comment	Critique of the post's content or Demagog's profile, often including accusations of bias or lack of objectivity.	„Serio nie powinniście tracić czasu na takie bzdury, russkij mir tworzy kilogramy fejków na minutę które zalewają polski internet". 4.03 "Seriously, you shouldn't waste time on such nonsense, the Russkij Mir creates kilograms of fakes per minute flooding Polish Internet." 4.03	Highlights controversial issues and often stimulates discussion - responses to such comments often clash with opposing views. People with one perspective, for example, those disagreeing with the criticism, can gain a sense of community in their opinion.

Comment Mentioning Another Person	Mentioning another person is a way to draw their attention to the content of the post, encouraging them to check it out.	„Damian patrz, nie potrzebny był nam demagog ale jest potwierdzenie pozdrów "zaufane zrodla" ;)))" 3.03 "Damian, look, we didn't need Demagog, but there's confirmation for 'trusted sources' ;)))" 3.03	Allows reaching people who may not have noticed the post or are not following the profile themselves. Through mentions, users are more likely to engage in discussion, respond, or at least react to the post. Mentions can be a source of new page likes once the mentioned person checks it out, thus expanding the community.
Comment with a Suggested Topic for Verification	Not directly related to the post's content but suggests another topic or specific fake news that, according to the commenter, Demagog should address	„Czy to że rok 2022 we Lwowie jest rokiem UPA to jest sprawdzona informacja,jeśli można ? Bo nie wiem czy na całej Ukrainie,ktoś gdzieś pisał,ale tego nie mogę nigdzie sprawdzić ?" 24.02 "Is it true that 2022 in Lviv is the year of UPA? Can you check it? Because I don't know if someone, somewhere wrote about it, but I can't verify it anywhere?" 24.02	Expresses uncertainty about the proposed topic and, at the same time, trust in the profile's accuracy and knowledge. Serves as a clear tool for social fact-checking by involving in expanding the content on the profile.
Comment Responding to the Post's Content	Directly refers in the comment to the content of the post but without clear approval or criticism of the portal's activities or criticism.	„Jest gdzieś nagranie gdy ta rakieta uderza. Trajektoria lotu nawet jest inna (powietrze ziemia niż ziemia-powietrze)" 27.02 "Is there a recording somewhere when this rocket had hit? Even the flight trajectory is different (air-to-ground than ground-to-air)." 27.02	Engaging in discussion or making a neutral comment often serves as an opportunity for further discussion in the responding comments. It often serves an informative but also persuasive function. Demonstrates an interest in the post's topic.
Comment Referring to Another Topic	Refers to another topic, often loosely related to the post's content. Sometimes includes elements of criticism. Different from the previous category, it doesn't necessarily propose a fact-checking topic - it may be more of a discussion starter.	„To jeszcze nic! Zobaczcie co jakaś antyszczepionkowa kretynka wkleiła pod postem o wojnie w Ukrainie. To jak oni mają sprane mózgi przechodzi jakiekolwiek pojęcie…" 4.03 "That's nothing! See what some anti-vax moron posted under the post about the war in Ukraine. It's beyond any understanding…" 4.03	Serves an informative function, introduces a new topic, shows an interest in fact-checking or broader news from the world.

Reply to Another Comment	A response from the profile's administrator or another user, a discussion with another comment or agreement with it, directly referring to the previous comment, e.g., by mentioning the person who posted it and clicking the "reply" option.	„Damian Ostrowski masz rację zwłaszcza, że Ty tak mówisz... i jeszcze Putin". 11.03. (odpowiedź na: „Kolejny artykuł typu: jest jedna prawda i to ta najprawdziwsza bo ja tak mówię".) "Damian Ostrowski, you're right, especially since you say so... and Putin too." 11.03.	Expresses interest in the topic, often showcases the user's knowledge. Multiple responses to comments are a sign of the dynamics of users' actions in social media. They can be emotional or ironic. Despite the clash of different views, they show critical thinking and a willingness to engage in a discussion.
Language Correction	Correction of the post's content, either linguistic or regarding typos.	„zestrzelono 2 czołgi" 24.02 "2 tanks were shot down." 24.02	Indicates that the user read the post's content and is engaged in its form, often more than in the content itself.
Link to a Charity Fundraising	In cases where the post's topic often appeared in the observed period, it is most often a link to charitable organizations, initiatives related to helping Ukraine.	„Tu info jak można pomóc https://m.facebook.com/story.php?story_fbid=10226942079363856&id=1409532436" 24.02 "Here's info on how to help	Indicates engagement in specific charitable actions. In the case of some posts related to charitable activities, they build a community around humanitarian values.
Technical Issue Comment	Draws the attention of administrators, for example, to a non-functional link in the post.	„Strona nie działa. Wszystko w porządku?" 24.02 "The page isn't working. Is everything okay?" 24.02	Indicates user engagement and the user's attempt to open the link to another webpage included in the post.

Table 1. Thematic categories of the posts and their functions.
Source: own work.

All the discusses types of comments are active actions, and although their significance for building a community varies, each of them signifies a willingness to interact with the profile administrators or other users. Through this communication, which reflects engagement in fact-checking activities—especially in the case of approving and critical comments, comments mentioning another person, and comments suggesting topics for verification—a community gathers around the profile with the potential to grow.

The posts that appeared on the profile during the study period can be classified into the following thematic categories: self-produced initiatives/products (e.g., podcasts, educational campaigns), war, health, climate and energy, foreign policy, changing profile pictures/cover photos, recruitment. The vast majority of posts were related to the war, with only 10 posts being associated with the "self-produced initiatives" category of the Demagog Association, 2 related to "health," 3 related to "foreign policy," 1 related to recruiting new team members, 1 related to energy issues, and all others were more or less loosely connected to the war-related theme.

The Demagog Association's Facebook profile has over 47,000 fans (as of June 2021) and has been active since March 2014. From the basic quantitative analysis, it is evident that during the study period, 149 posts were published, indicating a high level of activity by the administrators. This activity was particularly pronounced in the first part of the study period when the topic was fresh and elicited the most emotions. A total of 3,429 comments were posted under all the posts. The posts were shared 3,542 times, and there were 21,254 reactions to the posts. The detailed results of the quantitative analysis are presented in the table 2 below.

	Number of posts	Number of reactions	Number of comments	Number of shares	Topic with the highest numer of interactions
24.02	11	1801	163	487	War
25.02	7	750	133	55	War
26.02	2	138	9	33	War
27.02	8	3431	310	211	War
28.02	6	1130	334	342	War
1.03	4	706	108	127	War
2.03	5	652	88	212	War
3.03	11	1754	133	365	War
4.03	8	979	142	304	War
5.03	5	588	150	114	War
6.03	3	171	12	14	War
7.03	4	278	30	32	War
8.03	6	552	49	211	War
9.03	7	1249	208	203	War
10.03	8	1329	148	180	War
11.03	5	1234	764	122	War
12.03	2	151	35	19	War
13.03	2	169	15	22	War
14.03	4	249	57	26	War
15.03	4	591	103	118	War
16.03	4	434	41	40	War
17.03	5	262	88	20	War
18.03	4	575	131	85	War
19.03	3	189	10	24	War
20.03	2	213	11	15	War
21.03	4	340	18	31	War
22.03	4	315	51	39	Foreign policy
23.03	5	599	157	42	War
24.03	6	425	29	49	War

Table 2. The detailed results of the quantitative analysis.
Source: the author's research.

Posts with the highest cumulative number of reactions appeared on February 27th. This corresponds to the onset of the war when emotions among social media users were the most intensive. Additionally, on this day, a noteworthy total of 8 posts were published, representing one of the highest posting frequencies during the examined period. The trend of posting seems to decline over time, as the days with the most posts are concentrated in the first half of the research period. It should be noted that the post with the highest number of reactions is not simultaneously the one with the most comments;

it was added on February 28th. This may suggest that depending on the specific topic and format of the post, different levels of engagement are elicited. The same pattern holds for the number of shares, as the highest number occurred on the first day of the war (February 24th).

The topics that generated the most engagement during the study period were related to the war in Ukraine. This topic formed a community of interests and values and served as the foundation for the community's actions during that time. This was determined by a procedure that involved identifying the topic of the posts that received the highest number of interactions each day of the study month. It became evident that the topic of war consistently sparked the most interactions almost every day. Only on March 22nd, a post related to Marcin Błaszczak's statement on foreign policy generated more interactions than those related to the war.

Significantly, during the study period, various fact-checking organizations conducted numerous educational campaigns on misinformation. These initiatives may have attracted new followers to fact-checking profiles. Infographics aimed at educating people about misinformation were also highly popular, such as the one on February 28th, which received 252 shares.

There was a noticeable variation in user engagement based on the days. Early in the study period, charitable actions stirred strong emotions, and posts related to them were shared by a large number of users. This could indicate that, on social media, communities focus not only on interests, current events, and fact-checking but also on humanitarian values and the pursuit of truth. An example of expressing such values is a comment on the post from March 2nd, which said, "Thank you! Stay true during these times, all the best," and received 6 reactions.

It should also be emphasized that some posts may have been shared more frequently and to a larger audience due to paid advertising. However, the author does not have access to the advertising account of the profile and cannot determine which posts were sponsored.

The creators of the profile utilize various tools available on the Facebook platform to maximize their reach. The primary form of posts consists of graphics that clearly indicate the verification of information with labels such as "truth," "fake news," "manipulation," and "photomontage". These labels stand out against a uniform navy background. Other forms of content include links to articles that provide in-depth coverage of various social, economic, and other issues, short video clips, and posts promoting other communication channels of the Association, such as podcasts and newsletters. By using these various communication forms, the content can reach groups with diverse expectations regarding the format.

One form of community engagement in fact-checking activities is through reactions. Facebook provides several reaction options, including Like, Love, Haha, Wow, Sad, and Angry. The most common reaction is "Like," which suggests agreement with the content and serves as a form of praise for the post. These reactions are directed at both the content posted by the Demagog Association and comments from other users. Similarly, the "Angry" reaction is used in response to information that elicits negative emotions or is perceived as untrue. This is where the theory of social proof comes into play, which posits that in times of uncertainty, people tend to conform to the majority's beliefs and actions (Cialdini, 2013).

In addition to the previously discussed aspects of social fact-checking, it is also important to consider other forms existing on various social media platforms. While a comprehensive analysis of these forms goes beyond the scope of this article, they provide a broader context for the research.

One significant aspect of fact-checking is the practice of comments that highlight inaccuracies beneath posts from news outlets. When these comments receive a high

number of likes, algorithms tend to promote them by sorting them at the top of the comments section, giving them more visibility.

The role of influencers is yet another manifestation of building a fact-checking community. Content created by influencers often supports campaigns related to critical thinking and fact-checking. Education, through virtual opinion leaders, in line with the two-step flow of communication model, can influence the audience (Katz, Lazarsfeld, 1955). Communities built around influencers consist of engaged followers who actively participate in online discussions, leave comments, and share content. An example is the involvement of YouTuber Kasia Gandor in fact-checking initiatives related to the Demagog portal discussed above.

Can internet memes be considered a form of social fact-checking? On one hand, when memes playfully and ironically expose myths and conspiracy theories, they can stimulate critical thinking, and their virality is particularly strong. On the other hand, misleading memes may become more popular due to their sensational and emotional content. Short videos and graphics shared on platforms like Instagram and TikTok provide an opportunity to reach younger generations of users and shape their critical thinking. However, it is essential in this context to ensure that these memes do not become yet another source of fake news, embedding false information in the minds of the audience.

It is also worth noting the mechanisms of the social media platforms themselves. Facebook, for instance, has partnered with Agence France Presse to launch a fact-checking program that assesses the accuracy and reliability of published content. When content is deemed false, it is demoted in users' News Feeds, limiting its distribution. Additionally, pages that regularly share false information may lose access to advertising and monetization tools, and their reach could be reduced by up to 80 % (Szczęsny, 2019). These measures aim to reduce the spread of misinformation on the platform.

These various forms of social fact-checking contribute to the broader landscape of efforts to combat misinformation and promote critical thinking in the age of social media.

Summary and Conclusions

One of the threats is the question of the independence of fact-checking platforms, which has been occasionally raised in critical comments. Another issue is users relying on social proof, where in moments of uncertainty, simplified thinking prevails, and what the majority of people in their surroundings believe is considered correct behavior (Cialdini, 2013).

Another threat is that the Internet itself is a breeding ground for post-truth phenomena, and their scale may still surpass the efforts of some parts of the community. The sensationalism and emotional appeal of fake news easily capture the attention of a large audience, facilitating their uncontrolled and anonymous dissemination (Mroczka, 2022, pp. 86-128).

So, can social fact-checking have an impact on traditional media journalists to counter the spread of false information? Can such actions exert pressure on politicians to refrain from manipulating facts in their public statements? Although this is an open question and leaves room for further research, it should be considered that broad societal education through campaigns, such as those conducted by the Demagog Association, may yield the aforementioned results, and building a community of critical thinkers has the potential to counteract fake news through content sharing among profile authors, peers and acquaintances.

New media is both an opportunity and a threat. On the one hand, untrue content is spreading virally, and the problems of fake news or information bubbles are not the only challenges faced by the information society. On the other hand, in response to these

phenomena, initiatives appear to counter the spread of false information in the media.

The analysis of the Demagog Association's profile indicates further interesting research areas and provides a preliminary framework for understanding the community aspect of content verification organizations, offering a research tool in the form of comment categorization.

In response to the main research question, "what are the ways in which the social character of fact-checking is implemented on Facebook," it can be observed that the primary means are the active engagements of users on the profile. These activities hold particular significance for the formation of a community around the issue of content verification—they reflect users' willingness to interact with profile administrators or other users and often include suggestions for additional content to be fact-checked. During the studied period, the majority of interactions occurred under posts related to the war, indicating not only the significance and emotional impact of this topic but also the prevalence of misinformation surrounding it.

In the face of the war, numerous new initiatives have emerged with the aim of combating misinformation. Engaging in such actions provides social media users with an opportunity to participate in social fact-checking, which, in turn, has significant potential to educate and activate society in the field of digital literacy, including skills related to information verification. In the post-truth era, this form of education is of considerable importance and offers a chance to enhance information literacy among users of not only new, but also traditional media.

Well-organized fact-checking activities can serve as a catalyst for triggering mechanisms of community-based content verification through comments and reactions, not only on the organization's profile but also on news platforms. Additionally, the community vigilantly monitors the activities of influencers engaged in promoting the fight against misinformation, readily sharing the content they create as a part of social campaigns.

New media provide an opportunity for user engagement in verifying the accuracy of information, even in seemingly lighthearted forms such as memes or short videos. Furthermore, social media platforms themselves are getting involved in the battle against misinformation by implementing mechanisms that restrict the publication of certain content. Although the objectivity of these mechanisms is widely debated across various communities, it is unquestionably essential in the post-truth era to capitalize on every opportunity to promote the importance of truth and content reliability.

References

AGHAEI, S., NEMATBAKHSH, M. A., & FARSANI, H. K. (2012). Evolution of the world wide web: From WEB 1.0 TO WEB 4.0. *International Journal of Web & Semantic Technology, 3*(1), 1-10.

ALLCOTT, H., & GENTZKOW, M. (2017). *Social Media and Fake News in the 2016 Election.* Cambridge: National Bureau of Economic Research.

ALLCOTT, H., BRAGHIERI, L., EICHMEYER, S., & GENTZKOW, M. (2020). The Welfare Effects of Social Media. *American Economic Review, 110*(3), 629-676.

ARMSTRONG, A., & HAGEL, J. (1996). The real value of on-line communities. *Harvard Business Review, 74*, 134-141.

BOCHENEK, M., & LANGE, R. (2019). *Bezpieczne wybory. Badanie opinii o (dez)informacji w sieci.* NASK Państwowy Instytut Badawczy.

BONEK, T., & SMAGA, M. (2013). *Biznes na Facebooku i nie tylko. Praktyczny poradnik o promocji w mediach społecznościowych.* Warszawa: Oficyna a Wolters Kluwer business.

CHOMIAK-ORSA, I., & SMOLĄG, K. (2021). Wykorzystanie rozwiązań web w zarządzaniu wartością klienta w sektorze MSP. *Przegląd Organizacji, 12*(983), 24-31.

CIALDINI, R. (2013). *Wywieranie wpływu na ludzi. Teoria i praktyka.* Gdańskie Wydawnictwo Psychologiczne.

Definition of post-truth adjective from the Oxford Advanced Learner's. https://www.oxfordlearnersdictionaries.

com/definition/english/post-truth (online: November 10, 2021)

Demagog, Metodologia. https://demagog.org.pl/nasza-metodologia-jak-weryfikujemy-wypowiedzi (online: November 10, 2021).

DRYJAŃSKA, E. (2009). Social media jako narzędzie w rękach społeczeństwa obywatelskiego na przykładzie organizacji „Avaaz". *Global Media Journal−Polish Edition, 1*(5).

Dz.U.2020.0.805 t.j. - *Ustawa z dnia 29 grudnia 1992 r. o radiofonii i telewizji.*

ELLIOT, D., & CULVER, C. (1992). Defining and Analyzing Journalistic Deception. *Journal of Mass Media Ethics, 7*, 69-84.

GANS, H. (2004). *Deciding What's News: A Study of CBS Evening News, NBC Nightly News, Newsweek and Time.* Illinois: Northwestern University Press.

GILLIN, J. (2017). Fact-checking Fake News Reveals How Hard It Is To Kill Pervasive "nasty Weed". http://www.politifact.com/punditfact/article/2017/jan/27/fact-checking-fake-news-reveals-how-hard-it-kill-p/ (online: November 10, 2021)

GRUCHOŁA, M. (2020). Osoby starsze w świecie nowych mediów i technologii. *Teologia i Moralność, 15*(2), 28.

HARARI, Y. N. (2018). *21 lekcji na XXI wiek.* Kraków: Wydawnictwo Literackie.

IWANEK, J. (2006). Czwarta władza w demokracji współczesnej. *Władza, media, polityka.* Katowice: Wydawnictwo Uniwersytetu Śląskiego, 11-20.

KAPLAN, A. M., & HAENLEIN, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons, 53*(1), 59−68.

KAPOOR, K. K., TAMILMANI, K., RANA, N. P., PATIL, P., DWIVEDI, Y. K., & NERUR, S. (2018). Advances in social media research: past, present and future. *Informations System Frontiers, 20*(4), 531−558.

KATZ, E., & LAZARSFELD, P. F. (1955). *Personal Influence: the Part Played by People in the Flow of Mass Communication's.* New York: The Free Press.

KONOPKA, B. (2020). Szum informacyjny i jego rola w kształtowaniu warunków medialnych i kulturowych. *Transformacje, 1-2*(104-105), 167-187.

KRIPPENDORFF, K. (2004). *Content Analysis: An Introduction to Its Methodology.* London−New Delhi, 3.

KWIECIEŃ, M. (2017). "Fake news" słowem roku 2017 zespołu słowników Collinsa. Retrieved from http://www.slowanaczasie.uw.edu.pl/fake-news-slowem-roku-2017-zespolu-slownikow-collinsa/ (online: November 10, 2021).

LASSWELL, H. D. (2007). The Structure and Function of Communication in Society. In L. Bryson (Ed.), *The Communication of Ideas.* New York: 1948.

LISOWSKA-MAGDZIARZ, M. (2004). *Analiza zawartości mediów − przewodnik dla studentów.* Kraków.

ŁÓDZKI, B. (2017). "Fake news" - dezinformacja w mediach internetowych i formy jej zwalczania w przestrzeni międzynarodowej. *Polityka i społeczeństwo, 4*(15), https://www.ceeol.com/search/viewpdf?id=691152 (online: November 10, 2021).

MAZUR, J., & KUĆ, M. (2020). *Przestrzeń Społeczna, 1/2020* (19), 165-183.

MROCZKA, K. (2022). Fake newsy jako nowa kategoria zagrożenia systemu bezpieczeństwa ekonomicznego państwa w dobie kryzysu epidemicznego. *Przegląd Bezpieczeństwa Wewnętrznego, 26*(14), 2022.

MROZOWSKI, M. (2001). *Media masowe. Sztuka, rozrywka, biznes,* 122-124.

OLEKSY, P. (2022). Atak dezinformacyjny w polskiej cyberprzestrzeni. Instytut Europy Środkowej. Retrieved from https://ies.lublin.pl/komentarze/atak-dezinformacyjny-w-polskiej-cyberprzestrzeni/ (online: November 10, 2021).

PISAREK, W. (1983). *Analiza zawartości prasy.* Ośrodek Badań Prasoznawczych, Kraków.

POHL, A. (2018). Uprzedzenia w czasach postprawdy. In T. W. Grabowski, M. Lakomy, K. Oświecimski, A. Pohl (Ed.), *Postprawda. Spojrzenie krytyczne.* Kraków.

POPIELEC, D. U. (2014). Realizacja funkcji kontrolnej mediów na przykładzie dziennikarstwa śledczego. In I. Hoffman, D. Kępa-Figura (Ed.), *Wartości mediów t.2.*

POPIOŁEK, M. (2015). Serwisy społecznościowe w przestrzeni internetowej - (social)mediatyzacja życia codziennego, Kraków, *T. 58, nr 1* (221), 60−71.

POPIOŁEK, M. (2019). Bańka filtrująca i świadomość mechanizmów jej funkcjonowania wśród młodzieży. Wyniki badania przeprowadzonego wśród gimnazjalistów. *Zarządzanie Mediami, 7*(3), 159−171.

SKRZYPCZAK, J. (1999). *Popularna encyklopedia mass mediów.* Kurpisz S.A., 376.

SUCIU, P. (2022). Is Russia's Invasion Of Ukraine The First Social Media War? Retrieved from https://www.forbes.com/sites/petersuciu/2022/03/01/is-russias-invasion-of-ukraine-the-first-social-media-war/?sh=228bf-c271c5c (online: December 5, 2023).

SZCZĘSNY, J. (2019). Facebook wytacza wojnę fake newsom. W Polsce. Retrieved from https://antyweb.pl/facebook-polska-fake-news (online: December 5, 2023).