Research Article

When Checkers Mislead: A Study of (Dis)information Flows in Fact-Checking Practices in Poland

Roland Zarzycki[1]

1. Institute of Sociology, University of Wrocław, Poland

The main aim of the paper is to explore the dynamics and structure of (dis)information flows in order to better understand the disinformation phenomenon as such. We analyse a number of sociological models of (dis)information spread. We recognise key factors influencing these flows, including two static elements: education and ideological background, as well as two dynamic elements: censorship and fact-checking. As a result of this reflection, we conclude that explaining key aspects of the disinformation phenomenon requires going beyond these models and paying attention to meta-factors within a heuristic analysis. We critically reflect on the debate around the gains and losses of applying the aforementioned dynamic elements affecting the (dis)information flows and, as a result, social attitudes. On the basis of a content analysis of all published results of fact-checking processes carried out by leading Polish fact-checking platforms on their websites over a period of one month, we also provide the characteristics of fact-checking and its impact on the (dis)information flows in the Polish-speaking part of the internet. We show how the main Polish fact-checking portals operate in a biased manner, pursuing a specific political agenda. We provide a critical analysis of the social and political consequences of the diagnosed phenomena.

Corresponding author: Roland Zarzycki, roland.zarzycki@instytutnarutowicza.pl

1. Introduction: the concept

The information reality around us has become overwhelmed with disinformation. One might get the impression that in the 21st century we are facing unprecedented communication difficulties due to the blurring of the boundaries of truthfulness, while the rapidity of technological development and the hard-

to-understand flood of information make us vulnerable to these new challenges. However, although the term "fake news" has not been used for most of human history, this does not mean that the falsification of reality and the spread of rumours and false information have not been faced in the past. Moreover, for centuries we have been stuck with scientific theories and models of the world that later turned out to be false[1]. So, is it the emergence of "fake news" that means we find ourselves in some new circumstances? And if the circumstances are indeed new, what makes the difference?

Supposedly, Donald Trump posted 30,573 false or misleading claims on Twitter (now X) during his first four-year presidential term[2]. Interestingly, the heated political discussion around the disinformation spread by Trump does not credibly address the issue of how to precisely verify the truth or falsity of these 30,753 misleading claims. It seems that the polarisation between the parties is so extreme that the dispute itself is no longer rational or, in other words, it refers to two different structures of rationality that are mutually incompatible[3]. Here, we touch upon the problem of who has the right and competence to decide not about truthfulness, but about the very rules and the underlying concept of truthfulness.

In everyday life, the assertion of falsity is based on our beliefs about the nature of truth, embedded in our worldview, status and social roles. The success of the modern scientific method was that it developed schemes of thinking that made it possible to abstract to some extent from particular worldviews, statuses and social roles in our cognitive endeavours. Another component of this success was the testability of science and its openness to falsifiability. This, in turn, meant that, many times in contrast to the paths of intuitive or religious cognition, science accepts that no scientific model is ultimate and can claim absolute truth, for each model is subject to falsifiability.

One of our basic research assumptions is that in today's debate around disinformation, it is not the phenomenon of "fake news" that is unprecedented, but rather the attempt to apply science to the judgement of truth and falsity in a kind of absolute manner. From the perspective of the methodology of sciences, there is a fundamental difference between fallacious reasoning carried out within a given paradigm, where, as a result of an informed discussion, on the basis of certain principles of thought inherent to this given paradigm, we adjudicate fallacy within the paradigm, and a situation in which two incommensurable scientific paradigms clash[1]. In a broad political and ideological context, these two issues merge into one when certain actors in society attempt to settle certain issues in a definitive manner by abusing the authority of science. This absolute claim to objectivity encounters resistance at the same time, both within science itself and in the form of a critique of science, which points to a number of weaknesses undermining the scientific method, including its commodification and political

instrumentalisation. This phenomenon translates perfectly into the debate and practice of disinformation in the public space.

It is not the idea of this paper to discuss the condition of modern science. However, it is not possible to understand the dynamics and structure of the spread of disinformation as long as we ignore the public perception of these fundamental issues concerning the authority of science. In the Polish context, according to the Digital Poland[4] survey, only 11% of respondents agreed that scientists use reliable research methods in their research and another 42% tended to agree with this statement, which, in particular, gives very little more than half of the responses indicating trust in scientific methods of verifying truthfulness. At the same time, according to the same survey, 6% of people strongly believe that scientists conduct research at the behest of large corporations by which the results of this research are falsified, and a further 31% believe that this is rather the case. This means that in a significant number of cases the verification of information by scientific methods is not contested on the grounds of scientific polemics, but on the grounds of the unreliability of science as such, which in turn seems to be a consequence of the above-mentioned abuse of the authority of science to decide in an absolute manner on questions even though they are beyond the accepted cognitive paradigm.

2. Research goals and methodology

We analyse a number of existing models of (dis)information spread. Our aim is to explore the dynamics and structure of (dis)information flows in order to better understand the disinformation phenomenon as such. As a result of this reflection, we conclude that explaining key aspects of the disinformation phenomenon requires going beyond these models and paying attention to meta-factors within a heuristic analysis. And what are the typical factors explaining the spread of disinformation? Education, censorship, and fact-checking according to Jackson et al.[5]. This already shows that insufficient attention seems to be paid to at least the following meta-factors:

1. the ideological background and ideological convergence of the paradigms employed by the recipients of (dis)information;
2. the political and ideological conditioning of actions aimed at combating the spread of (dis)information, including so-called fact-checking;
3. the operational definition of disinformation itself, as (leaving aside academic disputes) fact-checking practice seems to pay little attention to the complexity of communication flows and thus

tends to judge the falsity or truthfulness of a given piece of information quite frivolously.

We recognised key factors influencing (dis)information flows, including two static elements: education (*ibidem*) and ideological background (including homophily), and two dynamic elements: censorship and fact-checking. To a certain extent we discuss the two static elements between the lines, but our research objectives focus on:

[conceptual aim] to critically reflect on the debate around the gains and losses of applying the aforementioned dynamic elements (censorship and fact-checking) affecting the (dis)information flows;

[empirical aim] to establish the characteristics of Polish fact-checking and its impact on the (dis)information flows in the Polish-speaking part of the internet.

We assign specific speculative research questions to the first objective:

[Q1.1] To what extent does the currently functioning method of creating models of the (dis)information flows take into account issues of political and ideological conditions, in particular within their discussion of the fundamentals of the methodology of sciences and the theory of truth?

[Q1.2] Is the overall effect of applying censorship and fact-checking to limit the spread of misinformation and disinformation positive?

Similarly, we formulate two specific empirical research questions related to the second objective:

[Q2.1] Do Polish fact-checking portals do their due diligence to function as objectively as possible?

[Q2.2] Do Polish fact-checking portals differentiate between methods of verifying misinformation and disinformation depending on the type of information being verified?

For questions Q1.1 and Q1.2, we carried out the research through critical reflection and discussion based on the available literature without formulating any detailed (empirical) hypotheses.

We associate two research hypotheses with questions Q2.1 and Q2.2, however:

[H2.1] The main Polish fact-checking portals operate in a biased manner, pursuing a specific political agenda.

[H2.2] Polish fact-checking portals do not adequately differentiate between methods of verifying fake news, abusing fact-checking methodologies as a result.

The methodology of the empirical part of the study includes an empirical research component: a content analysis of all published results of fact-checking processes carried out by leading Polish fact-checking platforms on their official websites over a period of one month (08.09.2023–07.10.2023). In order to

determine the leading Polish fact-checking platforms, the Digital Poland[4] report was employed, indicating a list of websites that describe themselves as fact-checking (in order depending on the popularity declared by respondents): oko.press (37%), konkret24 (35%), sprawdzimy.info (33%), fakenews.pl (32%), demagog.org.pl (23%), sprawdzam.afp.com (20%), pravda.org.pl (16%), fakehunter.pap.pl (12%) and odfejkuj.info (8%). Apart from the sites mentioned by the Digital Poland report, there are also other sites in Poland, such as Demaskator24, but due to their negligible market share, their findings have not been analysed.

Hypothesis H2.1 was verified by analysing the content and profile of the content verified by each of the leading Polish fact-checking platforms against their declarations on the one hand and the ethical standards for fact-checking platforms on the other. Hypothesis H2.2 was verified by analysing the types of content verified by each of the leading Polish fact-checking platforms against the UNHCR classification in comparison with the fact-checking method used.

3. Modelling the spread of fake news

The history of modelling riot behaviour as well as rumour diffusion is long. At its conceptual core lies the notion of a *threshold* that must be exceeded for a given message to become credible[6]. Thus, it follows that to a major extent the perception of a given message is determined by the number of people rather than the substance of the content or objectified criteria of truthfulness. At the same time, we know that the selection of messages that individuals disseminate and consider meaningful follows from a predetermined worldview[7]. Livan et al.[8] propose a learning model that explains how individuals are empowered to self-select narratives they want to be exposed to and how only a minority of them reject information that does not converge with their worldview. The problem, of course, is the unfounded assumption that this is a minority. This is consistent with the concept and findings of the so-called coherence theory of truth, and is also grounded in the fostering of polarisation and the creation of communication bubbles and echo chambers by the development of new media.

Many quantitative models for the spread and detection of fake news have been developed over a long time[9][10][11][12][13][14][15][16][17][18]. The mere integration of the know-how resulting from each of these hundreds of studies seems to be an important, necessary but also difficult research task that should be addressed at some point. From the perspective of this report, several models deserve special attention. An interesting, motivation analysis-oriented approach is proposed by Alwan et al.[19], analysing

multidimensionally the motivation to disseminate information by Internet users due to seven specified factors. The model developed by Aramaki et al.[20], on the other hand, takes into account the dynamics of the spread of fake news split into two segments: before and after a given news item has been recognised as fake. On the other hand, Havlin et al.[21] show that the capture of fake news is possible even before this moment, as the study of the flows shows structural differences in comparison with the flows of ordinary news.

An interesting classification of the phases of the spread of disinformation has also been proposed by Lynnette & Taeihagh[22]. Brady et al.[23] diagnose the phenomenon of moral contagion, which has a significant impact on the spread of fake news, and propose an explanatory MAD model. A key conclusion of these observations is to point out that group-identity-based motivations are crucial when communicating moral-emotional content, which in turn interacts with the very architecture of social media, which builds its power on these mechanisms. In this context, it is also worth mentioning the Centola et al.[24] study, which showed that in situations where individuals are isolated and exposed to communication with immediate neighbours with extremist views, highly questionable norms can emerge locally and be disseminated. An important approach to modelling fake news is presented in the work of Kopp et al.[25], where the traditional approach to modelling communication behaviour is enriched by capturing deception from a game theory perspective.

Fundamentally intriguing, however, is the fact that each of these models assumes that for a single message it is possible to determine whether it is false or not. Moreover, the neural networks used to construct these models rely on such elementary distinctions, as well as the selection of empirical material for their learning. This, in turn, would mean that it would be possible to determine clearly what, in the composition of such material, should be included and what should be excluded. Meanwhile, it seems that on a purely conceptual level, we do not have any universal practical tool for distinguishing real news from fake news. Here, the key distinction we mentioned in the previous section is whether we verify the falsity of a news item within a certain accepted paradigm or whether the falsity of a given news item should be considered in terms of competing paradigms. Unfortunately, both types of cases are mixed in communication practice. At the same time, the assumption that in the case of a single, atomic message it is possible to resolve its truthfulness, which is crucial in the construction of various paradigms, is highly debatable in practice.

Approaching the question of defining misinformation and disinformation from the perspective of pragmatic truth theory leads us to important observations. Among the 10 types of misinformation and

disinformation identified by UNHCR[26], there are three that are verifiable on the basis of technical content analysis (manipulated content, imposter content, sponsored content), four that are verifiable in a speculative-discursive manner on the basis of critical-retrospective discourse analysis (misleading content, false context, false connections, error), and three (fabricated content, satire and parody, propaganda), where verification is practically impossible without very strong additional assumptions concerning the concept of truth, the nature of communication, and often even metaphysical solutions (while it is also worth noting that this does not mean that all cases from the fabricated content category are difficult to resolve). Unfortunately, in the practice of public debate and within the currently popular fact-checking processes, no distinction is made here and often every type of misinformation and disinformation is treated as verifiable based on the prevailing sources.

As reliable sources point out[27], well-recognised phenomena such as confirmation bias, motivated reasoning or the availability heuristic can significantly affect the perception of facts. The definition of misinformation and disinformation becomes problematic in operational practice, and as a result, mathematical simplifications are used to create models of its spread, which, while allowing attractive models to be generated, unfortunately leave doubts about the adequacy of these models to the reality being modelled. For example, Brody et al.[28] define a model by assuming that fake news is any noise that disrupts information with a non-zero expected value of disruption. This framing of the issue is based on the important, though far from obvious, presupposition that we have a practical criterion for separating information from noise.

Some research teams construct their models in an attempt to capture the essence of this difficulty. As Cowen et al.[29] note, abstract models that work well for epidemiology are of limited use when the processing of signals through the network depends on the internal psychological structure or belief system of each vertex. In turn, as Baldassarri & Bearman[30] argue, models need to take into account the multidimensionality of the worldview.

At the same time, a large and rapidly growing area of research is the verification of the effectiveness of measures aimed at stopping the spread of misinformation and disinformation. As Patwa et al. note, "compared to traditional news media, the lack of effective regulation and fact-checking mechanisms over internet posts enables the creation and publication of false news to influence opinions and obtain low-cost revenues" (Patwa, et al., 2021). Thus, various types of internet platforms seem to provide a conducive environment for the propagation of false news[31]. Importantly, while flagging posts is not an effective

way to stop the spread of news, censoring them is already effective[32]. The problem is that, as it is non-transparent, it raises a number of important questions, including in terms of freedom of expression and the openness of science to critical positions. The solution, in a sense intermediate between flagging and censoring, seems to be fact-checking, which we will discuss in detail in the following sections.

Of particular note is the research of Acemoglu et al.[33], who show that homophily inhibits the wider circulation of articles but at the same time creates echo chambers within which reduced standards of reliability apply. Interestingly, as the authors point out, digital platforms themselves allow news with reduced credibility into spaces where we have greater polarisation, while more neutral and non-bullish news reaches a wider audience. In this sense, the spread of fake news is not solely determined by homophily but by an algorithm that relies on homophily to increase reach. This means that, from an economic perspective, the spread of misinformation and disinformation is aligned with the interests of the big digital platforms, which in turn opens up a fundamentally important political topic in the discussion.

In the context of the dynamic development of artificial intelligence, it is also worth mentioning that it is both an ally and an enemy in the fight against the spread of misinformation and disinformation: on the one hand, it can catch suspicious content, but on the other hand, it can also generate it. As Menczer & Yang[34] note, state-of-the-art LLM-content detectors cannot effectively distinguish between human and LLM-powered bots in the wild, which points to important dangers. Of course, an important question still remains as to how truth and falsehood are distinguished in the research material on which the neural networks are trained.

At the same time, increasing demands are being placed on digital platforms in the area of legislation. They are supposed to exercise due diligence and justify the measures used to combat disinformation on their services, as well as incorporate a range of fact-checking practices into their operations[35]. This creates a politically controversial situation in which a kind of censorship is enforced by the legislator, but the responsibility for restricting freedom of expression falls on the operators of the communication space.

4. Fact-checking industry: influence and bias

Fact-checking plays a noticeable and growing role these days. As Cherubini & Graves[36] already diagnosed in 2016, a plethora of fact-checking agencies have sprung up around the world. And while fact-

checking has always been around and similarly facts have always been verified, the 21st century has certainly brought at least a quantitative change in this regard. The flood of information has led to a need for rapid professionalisation in this industry (a positive phenomenon), but has also made it a rather lucrative and politically influential business (a potentially dangerous phenomenon). The quality and type of communication and the presence of fake news are indeed geopolitically conditioned (*ibidem*). Judging by our study and comparing it with the results of research conducted in the USA, one could hazard a guess to conclude that the quality of fact-checking actors is also geopolitically conditioned, although this is not the focus of our analysis.

Research by Cha et al.[37] shows that fact-checking affects young adults, but at the same time, a number of factors, such as a negative first impression, have a significant impact on the strength of this effect. As researchers point out, fact-checking platforms analyse many news items that are ambiguous and whose verification is questionable. A meta-analytic review conducted by Cohen et al.[38] showed that fact-checking has a large impact on political beliefs when truth is decided on a zero-one basis (d = 0.29), but this becomes much less clear when "truth scales" are used. The same study confirmed that the key importance for the perception of fake news and the results of its verification depends to a significant extent on pre-existing beliefs, ideology, and knowledge. This is directly related to the problem we pointed out in the first section, namely the lack of clear definitional boundaries in the case of certain types of disinformation, as well as the emergence of many intermediate cases. The perception of disinformation on the one hand and fact-checking reports on the other can be described using psychological mechanisms. And so, as Cha et al.[37] show, an experiment involving 11,145 young adults showed that these people are not convinced by ambiguous messages, which in turn is related to the widely recognised phenomenon of uncertainty-aversion. However, as the authors claim, fact-checking works well for relaxing extreme views.

A study by Porter and Wood[39] examined the effectiveness of fact-checking platforms. It was found that fact-checking significantly reduced beliefs in misinformation and persisted for several weeks. The authors estimated this impact at 0.59 points on a 5-point scale. However, as Tompkins[40] notes, fact-checking works, but only in quite specific circumstances. Citing Dan Kahan, he points out that people use false information if it confirms their position, but if it turns out to be false, they stop using it, which does not mean that they change their position. Another issue is whether a change in the state of knowledge results in a change in behaviour or voting preferences. As research by Barrera et al.[41] shows, it does not have to be this way. According to research by Adamic et al.[42] carried out on a group of over 10 million

American Facebook users, individual user choices were more important for limiting exposure to cross-cutting content than algorithmic ranking.

Aral et al.[43] showed that tweets containing falsehoods were 70% more likely to be retweeted than truthful tweets. Interestingly, this was not due, as it seemed at the beginning, to the fact that people spreading false tweets had more followers (because that is not the case), but tweets containing false information were simply more interesting and more novel, so, from a sociological point of view, could have been simply more useful in conversations. A similar study by Anderson et al.[44] showed that spreading fake news on Facebook is habitual and not a result of laziness or workload, which in turn is partly in contradiction with the research results of Pennycook & Rand[45]. It is worth noting that spreading the news is one thing and believing in its content is another.

In the large research project "Globalism Project 2021"[46], it was established that not all so-called conspiracy theories are equally popular and that their popularity is distributed very differently depending on the country[47]. Importantly, for example, in the case of the COVID-19 pandemic, there are great differences in the attitude towards the statement "coronavirus is a myth created by some powerful forces, and the virus does not really exist" and the statement "the truth about the harmful effects of vaccines is being deliberately hidden from the public". This also shows how important a role fact-checking platforms and the discussion about fake news promoted by prominent portals play in shaping the discourse around misinformation and disinformation. In this sense, fact-checking also influences the dynamics and structure of the trajectory of the spread of fake news. An important but under-researched problem is the biases of the fact-checking platforms themselves.

Levine et al.[48] conducted a study of the convergence of fact-checking platforms in the USA, focusing on the convergence and divergence of the selection of news to be checked and the results of this checking between The Washington Post and PolitiFact. As they pointed out, "the selection disagreement was indicated at the level of 22.6% [...] moderate agreement was observed for deceptiveness scaling and nearly complete agreement was observed for bottom-line attributed veracity"[48]. The researchers drew a number of important conclusions, in particular that the effect of fact-checking depends significantly on different fact-checkers and algorithmic bias. In a quite similar study, Nieminen & Sankari[49] analysed the fact-checking process used by PolitiFact, verifying a sample of 858 materials. Two problems diagnosed by the researchers are the unambiguous verification of complex propositions (33%) and the verification of statements whose truthfulness cannot be defined in practice, e.g., claims about the future and vague

claims (11%) (*ibidem*). Similar analysis was previously conducted between Fact Checker and PolitiFact, which found a high convergence in the case of those messages that were unambiguous, but considerable discrepancies in the ambiguous cases[50].

For example, as a side-result of our analysis of the mentioned 10 types of misinformation and disinformation distinguished by UNHCR[26], Polish fact-checking platforms are almost not interested in sponsored content (advertising or PR), even though it seems to constitute a significant share of misinformation and disinformation in general. We can speculate that this is due to fear of defamation lawsuits and confrontation with lawyers of large corporations. As Fraser notes, fact-checking itself is an element of the political game and the game of interests[51], which is also an issue that should remain in the centre of our attention.

5. Polish fact-checking platforms: an analysis

The study covered all fact-checking processes conducted and published by leading Polish fact-checking platforms in the period 08/09/2023–07/10/2023. A brief description of the investigated entities can be found below:

- oko.press: a political and social news website specialising in investigative journalism and fact-checking, operating since 2016 as part of the activities of the Citizen Control Centre "OKO" Foundation;
- konkret24: a political and social news website, which is a specialised unit of the Polish commercial television station TVN, owned since 2022 by Warner Bros. Discovery;
- sprawdzimy.info / odfejkuj.info (same website): was founded by the Cegielski Analysis Centre to verify and correct lies and manipulations; the portal was created with the support of the National Freedom Institute – Centre for the Development of Civil Society from the funds of the Programme for the Development of Civic Organisations for 2018-2030;
- fakenews.pl: an internet portal dealing with the broadly understood topic of fake news, run by the We Counter Disinformation foundation; the functioning of the website can be additionally characterised by highly intrusive messages asking for financial support;
- demagog.org.pl: a website run by the Demagog Association, the first Polish fact-checking organisation, which since 2014 has been verifying politicians' statements, checking election promises, fighting fake news and disinformation;

- checking.afp.com: a specialised unit of the French Agence France-Presse (AFP); the website focuses on verifying the accuracy of news from the socio-political area;
- pravda.org.pl: a website run by an association, which works to increase awareness of disinformation, improves the quality of public debate and is supported by the European Media and Information Fund;
- fakehunter: a social project for verifying content published on the internet, launched by the Polish Press Agency (a State Treasury company whose activities are based on the Act and has the character of a public press agency) together with GovTech Polska, whose initial goal was to unmask false news about the SARS-CoV-2; currently, the website is verifying requests sent by its users.

From the perspective of the research, an important issue is to verify whether the mentioned websites declare themselves as fact-checking and whether they declare the use of fact-checking standards. An analysis of the descriptions available on the websites indicates that they all declare themselves as fact-checking websites and, at least at the declaration level, they have the ambition to respect the standards. An important point of reference in this regard is the code developed by the European Fact-Checking Standards Network[52] as well as the code developed in 2016 by the International Fact-Checking Network (IFCN)[53]. At the same time, it is worth noting that phrases such as "From the set of all politicians' statements, we select those that are substantive and concern an objectively verifiable reality. The selection of the research sample is strictly qualitative and based on the criterion of material availability.", which can be found, for example, on the demagog.org.pl website, are highly speculative and ambiguous in nature.

As part of the study (analysis of published reports from fact-checking processes), it was noticed that these websites adopt very different approaches and various selections of news topics that are verified. The following significant features of the functioning of the platforms mentioned above were noted:

- oko.press – the specificity of the website, unlike most Polish fact-checking platforms, is that the materials posted there are (with exceptions) opinions and polemics with certain positions, rather than a verification of the truthfulness of any specific statements; in this sense, the fact-checking procedures used by oko.press are themselves speculative; one should also take into account that the very selection of content against which the polemic is presented has a strong right-wing bias (in the sense of stigmatising almost exclusively right-wing content) and thus, in terms of the selection of content itself, the portal should be recognised as biassed; at the same time, the importance of the

topics discussed is significant and wide-ranging; fact-checking frequency: 13 news items / month (low)

- konkret24 – a significant share of speculative materials is specific to the platform, while the result of verification should be recognised as manipulation ("misleading content") in many cases, which confirms speculative content; the selection of topics is evidently political (other types of disinformation are clearly marginalised), and the pro-liberal / anti-right-wing profile of the website is striking: among all verified fact-checking proceedings, over 36% had an anti-liberal tone and in 84% of these cases the verdict was false; at the same time, similarly, 36% had anti-right-wing connotations, and only 5% of cases were found to be false; of course, the website's editors may argue that the reason for this is the more frequent disinformation in right-wing messages, but this does not seem to correspond to reality, taking into account the findings of other websites; fact-checking frequency: 53 / month (significant)
- sprawdzimy.info and odfejkuj.info – the website did not post a single fact-checking result in the analysed period; fact-checking frequency 0 / month (the portal has probably suspended operations)
- fakenews.pl – the selection of materials is diverse, comes from various thematic areas and has no clear political orientation, which, however, is difficult to determine due to the small sample, as the portal operates with very low intensity; attention is drawn to the relatively unambiguous formulation of the verified content, which in turn creates the possibility of zero-one verification; at the same time, the selection of materials on this website seems controversial in another way: most of them seem insignificant, which in turn could be attributed to the avoidance of important and controversial topics; fact-checking frequency: 8 / month (very low)
- demagog.org.pl – the website presents fact-checking processes in two categories: "(fake) news" and "statements"; in the case of the statements section, politicians' statements are presented, which are usually quite unambiguous, which in turn allows for their unambiguous verification, and such verification (highly reliable compared to other websites) is provided by the Demagog website; at the same time, in this category, the portal does not avoid important and difficult topics; what is noteworthy is the political balance of the selection of news items which are being verified (the 35 statements included: 26% of statements by centre-liberal politicians, 31% by right-wing politicians, 31% by left-wing politicians and 11% by independents), while the assessment of the truthfulness of the selected statements may be controversial (11% false in the statements of centre-liberal politicians, 36% false in the statements of right-wing politicians; 18% false in the statements of left-wing politicians and 25% false in the statements of independent people) (it is worth noting that the set of

clearly false answers is complemented not only by true answers, because the portal also uses other categories, such as "inconclusive" or "partially false"); a negative impression about the reliability of the portal is created by the second category presented by the website ("(fake) news"), where several specific topics dominate, including primarily the issue of the (harmlessness) of 5G technology, as well as the issue of vaccinations; additionally, a significant part of the verified news comes from one specific portal, which bears the hallmarks of an ideological struggle; especially as much of the verified information seems completely insignificant; however, the most important reservation is the previously mentioned problem with the selection of the verification method, namely, news that questions the findings of the scientific mainstream is in many cases verified by referring to the findings of this mainstream science; fact-checking frequency: 97 / month (very high)

- checking.afp.com – the website can be characterised by reliability in verifying the presented news, based on well-argued sources, as well as the usage of technological analyses (on a scale unheard of on other platforms: 47%); the verified information is quite unambiguous, which allows for the formulation of unambiguous results of the investigations; the portal does not tend to discuss political topics; at the same time, it is worth noting that the selection of verified content is problematic because in many cases it is insignificant (the question arises: why verify it if many more important items are not verified?), but also because there is a clear dominance of news with anti-Ukrainian content, which the website verifies; eventually, the website's decision to publish only content that has been verified as false is questionable; fact-checking frequency: 15/month (low)
- pravda.org.pl – during the period under review, the website verified only anti-Ukrainian content and did so in an inaccurate and questionable manner; however, it is difficult to draw broader conclusions considering the very low activity of the website's editorial staff; fact-checking frequency: 4/month (very low)
- fakehunter – unlike other websites, Fakehunter verifies content sent to the editorial office by users with a request for verification; since it is difficult to determine how (if at all) the editors additionally filter these queries, based on the principle of presumption of innocence, it should be stated that Fakehunter selects content reliably; the queries formulated by users themselves are of a very diverse nature, but it is worth noting that a major part of them is utterly insignificant or even weird or trivial; in the vast majority of cases, the portal's editors present reliable arguments when verifying users' queries; fact-checking frequency: 78/month (high).

In the entire sample of 268 analysed messages, only less than 11% were classified as unambiguously true; in this context, a key question arises as to whether: (a) this is a structural phenomenon and only 11% of news in public circulation is clearly true; (b) for some reason, fact-checking websites tend to select false news; (c) for some reason, fact-checking websites' verification processes are asymmetrically tilted towards questioning the truthfulness of the analysed content.

Among the content analysed by Polish fact-checking portals, there was a vestigial (<1%) amount of misinformation and disinformation from the "sponsored content (advertising or PR)" category (according to UNHCR); similarly, questions arise as to whether: (a) this is a structural phenomenon and news of this type does not appear in public circulation; (b) for some reason, fact-checking websites do not analyse this type of content.

Some websites, such as Demagog, do analyse jokes ("satire and parody" in the UNHCR classification), rightly pointing out that they also contribute to disinformation; interestingly, however, their share in the total sample is about 2%, which seems to be an extremely low value, considering how many different types of memes flood the Internet; in other words, it does not seem that humorous and sarcastic materials account for only 2% of the disinformation, which in turn reveals a bias in the selection of categories in yet another way.

The analysis of the most popular topics undertaken by fact-checking websites in Poland in 2022, carried out by the Media Education Centre[54], showed that the selection of this news bears signs of bias; in particular, among the 20 most popular tweets, three topics were clearly visible: current Polish politics, COVID and issues related to refugees from Ukraine. Interestingly, most of these publications came from three websites: Demagog, Konkret24 and FakeNews.pl[54]; the question arises again whether this is due to structural factors or some kind of conscious selection asymmetry.

As a rule, the analysed websites do not distinguish between different types of misinformation and disinformation (in accordance with the adopted UNHCR classification), so in particular: (a) they verify news from the broadly understood category of conspiracy theories by referring to official communications from public institutions; (b) they verify the questioning of commonly accepted findings of mainstream science by referring to these commonly accepted findings of mainstream science.

Interestingly, only about 7% were messages verified through technical analysis (photomontage, deepfake, etc.).

In the general debate on disinformation, it is often indicated that the Russian Federation has a significant contribution to digital disinformation; the examined material does not provide grounds to confirm this state of affairs, because in a marginal number of cases, investigative journalists pointed to the Russian origin of the news verified as false.

Although it is difficult to assess the political or ideological importance of a given piece of news, it seems that fact-checking platforms intertwine the verification of completely trivial news with highly politically charged news; insignificant information is often verified, while very significant statements are not verified that thoroughly. Of course, this can be explained by the difficulty of verification, but the verdict "unresolved" or "partial truth" appears in the analysed reports, as it can be seen that consideration of information that is difficult to verify is practised; one can pose questions on the intentionality of such an approach.

6. Discussion and conclusions

The basic conclusion of the study is the political and ideological conditioning of the discourse on misinformation and disinformation, including so-called "fake news". Although it is very common nowadays to talk about the enormous scale of this phenomenon, determining in practice what it is exactly turns out to be highly problematic. It seems that an important voice in this discussion should be the voice of fact-checking platforms established to formulate expert insights. However, both the literature discussing this issue in the global context and empirical research carried out within the Polish-speaking part of the internet point to important limitations of such a perspective. Firstly, according to research by Digital Poland[4], few (5%) users only take into account the indications of fact-checking platforms; secondly, the reliability of the functioning of Polish platforms of this type is highly questionable, which is a direct conclusion from our study; and thirdly (which should perhaps be seen as a consequence of the first two points), the analysis of the selected case studies showed that the impact of these platforms on the spread of misinformation and disinformation is unnoticeable. They seem to convince the convinced.

If the area of misinformation and disinformation within the Polish-speaking internet is defined in accordance with the indications of fact-checking platforms, it should be considered a marginal phenomenon. However, in practice, most users operate differently and use different methods to decide whether various content is true or false. Among these, it is certainly worth mentioning techniques such as self-verification, opinions of online influencers, or other types of journalistic content, such as

television programmes. This is significant as there are no right-wing-oriented platforms on the Polish fact-checking scene that nominally declare themselves fact-checking. This is somewhat understandable, taking into account the ideals of the Enlightenment.

According to the presented analyses, at least some of the modern models illustrating the dynamics and structure of the (dis)information flows indicate the key importance of the ideological background and homophily (among static factors), the high effectiveness of content censorship and the limited effectiveness of fact-checking. This seems to lead to ideological and worldview tensions when verifying the truthfulness of fabricated content, satire and parody, and propaganda, and, consequently, to the abuse of the authority of science as such, harming its image.

As the latest research shows, limiting reach could be recognised as another effective method[5], which WhatsApp has been doing for some time. This is important because it is exactly a moment in which something nominally free starts being manipulated and/or regulated. This is also a place where financial issues (because outreach means profits) begin to intertwine with discourse control. At the same time, it seems that suppressing the discussion may even have the effect of catalysing the spread of fake news in the bubble of people who have been excluded from a given mainstream discourse. Interestingly, censorship is often the result of political pressure (cf. [55]), when politicians of certain formations try to cede the enforcement of certain attitudes in the communication sphere to business. What often draws attention is the ferocity with which alternative discourse is suppressed, many examples of which could be observed during the COVID-19 pandemic.

Despite the significant dangers associated with treating fact-checking as a source of indisputable facts signalled by Ireton & Posetti[27], the current political debate around fact-checking has turned towards censorship. It seems that this is largely a consequence of the COVID-19 pandemic, during which there were attempts, in an irrationally emotional way, to remove completely from the space of public discourse any content that could question the shape of the adopted anti-epidemic policies. Censorship, also as morally ambivalent, turns out to be an effective tool for suppressing misinformation and disinformation in a given digital space, which does not lead, however, to the disappearance of certain beliefs. On the other hand, it is worth noting that collective intelligence works better when mechanisms encourage confrontation between distinct viewpoints[56].

In light of the above results, the attitude of respondents expressed in the Digital Poland study[4] may be worrying. Forty-one per cent definitely considered "deleting or blocking social media accounts of people

/ organisations that most often publish false information (fake news)" as a good method of combating disinformation, and another 30% rather agreed with it, while 69% considered "automatic deleting false information by internet giants such as Facebook, Instagram, Google after 1 hour from publication on social media". These results indicate a high level of unawareness of structural violence and its consequences among Polish society.

A study of the functioning of Polish fact-checking platforms indicated an array of discrepancies between the practices used by these platforms and the standards applicable in the industry. The image of Polish fact-checking emerging from the study indicates the political and ideological involvement of the editorial staff of these websites, which is far from the expectations regarding their objectivity. It is revealed on several levels, including: in the selection of verified content from a political angle, in the importance of verified content, in the omission of a number of significant topics, and in the methods used for verification. We can speculate as to the intentionality of these activities, which may be unconscious but may also be aimed at increasing the commercial attractiveness of websites or implementing a specific ideological agenda. In this sense, the validity of the adopted hypothesis H2.1 depends on resolving the issue of intention. Without falling into psychologism, however, it can be said that the activity of Polish fact-checking platforms indicates their highly counter-objective attitude.

Hypothesis H2.2 was confirmed in the study because all types of news were verified by leading websites using very similar methods, in particular by referring to the mainstream message and official positions of the institutions, even in circumstances where the given news questioned the credibility of this mainstream message. At the same time, there are known cases of situations in which the mainstream message turned out to be completely wrong[57].

It seems that this fact-checking bias is recognised by internet users and treated as a vehicle of ideological influence in the context of discourse wars. Mixing verification of trivial and trivially erroneous news with news of great complexity and importance, using similar methods, further obscures the image of fact-checking. Due to this unreliable image and suspicions of instrumental use, people with different worldviews and/or rationalities, in particular, are not willing to accept the results of investigations as an important point of reference. The general level of trust in fact-checking platforms in Poland should be considered low. It also seems that this lack of trust indirectly affects science itself, to the authority of which fact-checking platforms very often refer. An important question is the impact of the wide dissemination of models based on self-learning networks (AI) on this sector and the condition of the

debate about truth. Of course, due to the limited scope of this paper, many issues could not be clarified sufficiently. This should be considered a good critical starting point for further research.

About the Author

Dr Dr Roland Zarzycki – affiliated with the Institute of Sociology, University of Wroclaw, PhD in mathematics, PhD in humanities. Also the Deputy Director at the Gabriel Narutowicz Institute of Political Thought. Researcher, academic teacher, speaker at international conferences, and research and social projects manager and coordinator. Collaborates internationally on political and economic issues. Co-authored and coordinated multiple social, cultural, and educational projects. Interested in social and political economy, critical political thinking and the role of fantasy and ideology in everyday life.

Statements and Declarations

Funding

This article was created based on research conducted within the framework of the Central European Digital Media Observatory project financed by the European Union under the call: DIGITAL-2023-DEPLOY-04, project 101158609.

Potential Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Ethics

No human subjects were involved in the empirical research.

Data Availability

All websites used as empirical material for this study remain accessible online, with a detailed description of the analysed sites provided in the methodological section of this article.

References

1. a, b*Kuhn ThS (2012). The Structure of Scientific Revolutions. 50th ed. Chicago, IL: University of Chicago Pres s.*

2. ^Kessler G (2021). "Trump Made 30,573 False or Misleading Claims as President. Nearly Half Came in His Final Year." The Washington Post. bit.ly/46PQcyz.
3. ^McIntyre A (1988). Whose Justice? Which Rationality?. University of Notre Dame Press.
4. a, b, c, dDigital Poland (2022). "Dezinformacja Oczami Polaków" [Disinformation Through the Eyes of Poles]. Digital Poland FOUNDATION.
5. a, bJackson MO, Malladi S, McAdams D (2022). "Learning Through the Grapevine and the Impact of the Breadth and Depth of Social Networks." **119**(34). doi:10.1073/pnas.2205549119.
6. ^Granovetter M (1978). "Threshold Models of Collective Behavior." Am J Sociol. **83**(6):1420–1443.
7. ^Hahn KS, Iyengar S (2009). "Red Media, Blue Media: Evidence of Ideological Selectivity in Media Use Get Access Arrow." J Commun. **59**(1):19–39. doi:10.1111/j.1460-2466.2008.01402.x.
8. ^Livan G, Sikder O, Smith RE, Vivo P (2020). "A Minimalistic Model of Bias, Polarisation and Misinformation in Social Networks." Sci Rep. **10**:5493. doi:10.1038/s41598-020-62085-w.
9. ^Africano Franco DR, Barragán Moreno S, Carrillo Barbosa RL, Guzmán Rincón A, Rodríguez-Canovas B (2023). "Social Networks, Disinformation and Diplomacy: A Dynamic Model for a Current Problem." Humanit Soc Sci Commun. **10**:505. doi:10.1057/s41599-023-01998-z.
10. ^Apostolidis L, Kompatsiaris I, Makedas T, Papadopoulos S, Papadopoulou O, Poldi F (2022). "MeVer NetworkX: Network Analysis and Visualisation for Tracing Disinformation." Future Internet. **14**(5):147. doi:10.3390/fi14050147.
11. ^Arini HM, Liu X, Mohawesh R, Wu Y, Yin H (2023). "Semantic Graph Based Topic Modelling Framework for Multilingual Fake News Detection." AI Open. **4**:33–41.
12. ^Batistela CM, Piqueira JRC, Zilbovicius M (2020). "Daley-Kendal Models in Fake-News Scenario." Physica A Stat Mech Appl. **548**:123406.
13. ^Franceschi J, Pareschi L (2022). "Spreading of Fake News, Competence and Learning: Kinetic Modelling and Numerical Approximation." Phil Trans R Soc A. **380**:20210159. doi:10.1098/rsta.2021.0159.
14. ^Franceschi J, Pareschi L, Zanella M (2022). "From Agent-Based Models to the Macroscopic Description of Fake-News Spread: The Role of Competence in Data-Driven Applications." SN Partial Differ Equ Appl. **3**(6):68. doi:10.1007/s42985-022-00194-z.
15. ^Liming W, Pengwei Z, Yang W, Yunjian Z, Zhen X (2021). "Multimodal Fusion with Co-Attention Networks for Fake News Detection." In: Findings of the Association for Computational Linguistics. ACL-IJCNLP. p. 2560–2569.

16. ^Qian Y, Ye Q, Yuan H, Zhang Y, Zheng J (2021). "Improving Fake News Detection with Domain-Adversarial and Graph-Attention Neural Network." Decis Support Syst. **151**:113633.

17. ^Ren Y, Zhang J (2021). "Fake News Detection on News-Oriented Heterogeneous Information Networks Thr ough Hierarchical Graph Attention." In: International Joint Conference on Neural Networks (IJCNN). IEEE. p. 1–8.

18. ^Aldini A (2022). "On the Modelling and Verification of the Spread of Fake News, Algebraically." J Logic Co mput. **32**(6):1272–1291. doi:10.1093/logcom/exac015.

19. ^Alwan A, Garcia EP, Kirakosian AT, Weiss AP (2021). "Toward a Comprehensive Model of Fake News: A Ne w Approach to Examine the Creation and Sharing of False Information." Societies. **11**(3):82. doi:10.3390/soc1 1030082.

20. ^Aramaki E, Kobayashi R, Murayama T, Wakamiya S (2021). "Modelling the Spread of Fake News on Twitte r." PLoS ONE. **16**(4):e0250419. doi:10.1371/journal.pone.0250419.

21. ^Havlin S, Levy O, Li D, Sano Y, Takayasu H, Takayasu M, Wu J, Zhao J, Zhao Z (2020). "Fake News Propagat es Differently From Real News Even at Early Stages of Spreading." EPJ Data Sci. **9**:7. doi:10.1140/epjds/s13688 -020-00224-z.

22. ^Lynnette HXN, Taeihagh A (2021). "How Does Fake News Spread? Understanding Pathways of Disinforma tion Spread Through APIs." Policy Internet. **13**:560–585.

23. ^Brady WJ, Crockett MJ, Van Bavel JJ (2020). "The MAD Model of Moral Contagion: The Role of Motivation, Attention, and Design in the Spread of Moralized Content Online." Perspect Psychol Sci. **15**(4):978–1010. doi: 10.1177/1745691620917336.

24. ^Centola D, Macy M, Willer R (2005). "The Emperor's Dilemma: A Computational Model of Self-Enforcing N orms." Am J Sociol. **110**(4).

25. ^Kopp C, Korb KB, Mills BI (2018). "Information-Theoretic Models of Deception: Modelling Cooperation and Diffusion in Populations Exposed to 'Fake News'." PLoS ONE. **13**(11):e0207383. doi:10.1371/journal.pone.0207 383.

26. a, bUnhcr (2021). "Using Social Media in Community Based Protection: A Guide." Office of the United Nation s High Commissioner for Refugees.

27. a, bIreton C, Posetti J (eds) (2018). "Journalism, 'Fake News' & Disinformation. Handbook for Journalism Edu cation and Training." United Nations Educational, Scientific and Cultural Organization.

28. ^Brody D, Hughston L, Macrina A (2022). "Modelling Information Flows in Financial Markets." In: Brody D, Hughston L, Macrina A. Financial Informatics: An Information-Based Approach to Asset Pricing. World Scie

ntific Publishing.

29. ^*Cowen L, De Ruiter J-P, Rabb N, Scheutz M (2022). "Cognitive Cascades: How to Model (and Potentially Co unter) the Spread of Fake News." PLoS ONE. **17**(1):e0261811. doi:10.1371/journal.pone.0261811.*
30. ^*Baldassarri D, Bearman P (2007). "Dynamics of Political Polarization." Am Sociol Rev. **72**(5). doi:10.1177/0 00312240707200507.*
31. ^*Ghorbani AA, Zhang X (2020). "An Overview of Online Fake News: Characterisation, Detection, and Discus sion." Inf Process Manag. **57**:102025.*
32. ^*Clark N, Ko S, Sai Z (2022). "Modelling the Spread of Misinformation on Social Media." YT. youtube.com/w atch?app=desktop&v=8b258RJR1TM.*
33. ^*Acemoglu D, Ozdaglar A, Siderius J (2021). "A Model of Online Misinformation." NBER Working Paper. No. 28884. nber.org/papers/w28884.*
34. ^*Menczer F, Yang K-CH (2023). "Anatomy of an AI-Powered Malicious Social Botnet." Observatory on Socia l Media Indiana University. arXiv: 2307.16336v1.*
35. ^*Miguel R (2023). "Platforms' Policies On Ai-Manipulated And Generated Misinformation." EU DisinfoLab.*
36. ^*Cherubini F, Graves L (2016). "The Rise of Fact-Checking Sites in Europe." Reuters Institute for the Study of Journalism.*
37. a, b*Cha M, Kang J-H, Park JY, Park S (2021). "The Presence of Unexpected Biases in Online Fact-Checking." H arvard Kennedy School Misinformation. **2**(1). doi:10.37016/mr-2020-53.*
38. ^*Cohen J, Holbert L, Morag Y, Walter N (2020). "Fact-Checking: A Meta-Analysis of What Works and for Wh om." Polit Commun. **37**(3):350–375. doi:10.1080/10584609.2019.1668894.*
39. ^*Porter E, Wood TJ (2021). "The Global Effectiveness of Fact-Checking: Evidence From Simultaneous Experi ments in Argentina, Nigeria, South Africa, and the United Kingdom." Proc Natl Acad Sci USA. **118**(37). doi:10. 1073/pnas.2104235118.*
40. ^*Tompkins A (2020). "Is Fact-Checking Effective? A Critical Review of What Works — and What Doesn't." D W Akademie. p.dw.com/p/3joav.*
41. ^*Barrera O, Guriev S, Henry E, Zhuravskaya E (2020). "Facts, Alternative Facts, and Fact Checking in Times of Post-Truth Politics." J Public Econ. **182**:104123.*
42. ^*Adamica LA, Bakshy E, Messing S (2015). "Exposure to Ideologically Diverse News and Opinion on Facebo ok." Science. **348**(6239):1130–1132. doi:10.1126/science.aaa116.*
43. ^*Aral S, Roy D, Vosoughi S (2018). "The Spread of True and False News Online." **359**(6380):1146–1151. doi:10.1 126/science.aap95.*

44. ^Anderson IA, Ceylana G, Wood W (2021). "Sharing of Misinformation is Habitual, Not Just Lazy or Biased." **120**(4).

45. ^Pennycook G, Rand DG (2019). "Lazy, Not Biased: Susceptibility to Partisan Fake News is Better Explained by Lack of Reasoning Than by Motivated Reasoning." Cognition. **188**:39–50. doi:10.1016/j.cognition.2018.06.011.

46. ^Rogers De Waal J (2021). "Globalism Project 2021: Highlights & Methodology." YouGov. bit.ly/46wNT3p.

47. ^Kirk I (2022). "What Conspiracy Theories Did People Around the World Believe in 2021?" YouGov. bit.ly/46vrovW.

48. a, bLevine TR, Markowitz DM, Moore AD, Serota KB (2023). "Cross-Checking Journalistic Fact-Checkers: The Role of Sampling and Scaling in Interpreting False and Misleading Statements." PLoS One. **18**(7):e0289004. doi:10.1371/journal.pone.0289004.

49. ^Nieminen S, Sankari V (2021). "Checking PolitiFact's Fact-Checks." J Stud. **22**(3):358–378. doi:10.1080/1461670X.2021.1873818.

50. ^Lim CH (2018). "Checking How Fact-Checkers Check." doi:10.1177/20531680187868.

51. ^Fraser M (2023). "Czy Facebook Wykorzystuje Fact-Checking do Własnej Gry Politycznej?" [Does Facebook Use Fact-Checking for Its Own Political Game?]. CyberDefense24.pl. bit.ly/3tqgpVO.

52. ^EFCSN (2022). "The European Code of Standards for Independent Fact-Checking Organisations." European Fact-Checking Standards Network. efcsn.com/code-of-standards.

53. ^Babraj R (2019). "Czym Jest Fact-Checking? – Zarys Inicjatyw na Świecie i w Polsce" [What is Fact-Checking? – An Outline of Initiatives in the World and in Poland]. Nask Cyber Policy. bit.ly/3tvRwbw.

54. a, bCEM (2023). "Polityka i Ukraina: O Czym Pisały Polskie Serwisy Fact-Checkingowe w 2022 Roku?" [Politics and Ukraine: What Polish Fact-Checking Services Wrote About in 2022?]. Centrum Edukacji Medialnej. bit.ly/48WqOJ2.

55. ^Money (2023). "Wojna w Izraelu to Problem dla Elona Muska. Platformie X Grozi Całkowite Wyłączenie" [War in Israel is a Problem for Elon Musk. Platform X is Threatened with Complete Shutdown]. money.pl. bit.ly/46qLnvH.

56. ^Greenstein S, Gu G, Zhu F (2020). "Ideology and Composition Among an Online Crowd: Evidence From Wikipedians." Manag Sci. **67**(5):3067–3086. doi:10.1287/mnsc.2020.3661.

57. ^Healio (2009). "Cigarettes Were Once 'Physician' Tested, Approved." Hemonc Today. bit.ly/3S6ucLw.

Declarations

Funding: This article was created based on research conducted within the framework of the Central European Digital Media Observatory project financed by the European Union under the call: DIGITAL-2023-DEPLOY-04, project 101158609.

Potential competing interests: No potential competing interests to declare.