The (Possible) Use of AI Tools for Processing Texts in Journalism in Bulgarian

Ruslana Margova
Big Data for Smart Society Institute
(GATE)
Sofia University
5, James Boucher St., Sofia, 1164,
Bulgaria
ruslana.margova@gate-ai.eu

This study examines the technological gaps in text-processing AI tools available to Bulgarian-language journalists and how these tools might better support journalistic practices. Through a systematic analysis of current technologies across three key domains—monitoring and information gathering, content production, and content dissemination—the research reveals significant disparities between international standards and local capabilities. While some resources exist for Bulgarian journalism, including news aggregators, translation services, these tools often lack transparency, update infrequently, or provide insufficient functionality for professional journalistic needs. Large language models (LLMs) offer promising possibilities but remain underutilised in Bulgarian newsrooms. The article provides a case study about the practical use of AI. The study recommends strategic investment in language-specific AI, targeted training, transparency standards, and ethical frameworks to improve journalistic capacity and information quality in Bulgaria, as trustworthy journalism must reach wider audiences to drown out disinformation.

Keywords: *artificial intelligence, Bulgarian, journalism, text processing, computational linguistics*

1. Introduction

Journalism operates under severe time pressures, demanding quick decision-making and substantial output from professionals who often rely on intuition honed through experience. The recent rise of generative artificial intelligence (AI) has intensified debates around misinformation proliferation and the potential displacement of human journalists. Despite these emerging challenges, journalism remains an inherently human-centred field, though its practices continue to evolve with technological advances.

This study investigates a critical research question: What technological gaps exist in text-processing tools available to Bulgarian-language journalists, and how might these tools be enhanced to better support journalistic needs in content creation and quality reporting? Our analysis concentrates on current text-processing technologies in Bulgarian, deliberately

https://doi.org/10.47810/JCLIB.1.2025.05

excluding OSINT[1] and visual tools to maintain focus, while examining AI's dual role as both collaborative assistant and subject of ethical consideration.

The research examines three primary applications of AI in journalistic practice:

1. Monitoring and information gathering
2. Content production
3. Dissemination

Our simple framework for evaluating AI tools for journalism includes language support, transparency and usability of the tools in Bulgarian. Additionally, we conduct a discourse analysis of selected articles where artificial intelligence is explicitly acknowledged as a co-author of journalistic content, exploring the implications of this emerging practice and raising pertinent ethical questions.

Identifying and addressing technological gaps faced by Bulgarian journalists is crucial for enhancing newsroom efficiency, accuracy, and adaptability in today's rapidly evolving digital media landscape. Many journalists currently lack access to advanced technologies due to financial constraints, insufficient training opportunities, or organizational resistance to change – factors that significantly impair their capacity to gather, verify, and present information effectively. By first addressing text-processing AI tools (visual applications to be explored in future research), this study aims to provide a foundation for Bulgarian journalists to leverage innovative technologies for data analysis, data gathering and content creation. Such technological integration ultimately enhances journalistic quality and credibility and would lead to the sustainability of the Bulgarian media industry.

Drawing upon years of journalistic experience, the author uses his own experience, also incorporating insights gleaned from informal discussions with fellow journalists. This blend of direct professional involvement and peer perspectives provides a nuanced understanding of the issues at hand within the journalistic community.

2. Context and Related Work

Research on technological integration in journalism has evolved significantly over time. In 2017, the first comprehensive survey on newsroom technologies revealed a concerning lack of technical expertise, with merely 5% of newsroom staff possessing technology-related degrees, only 2% of newsrooms employing technologists, and a mere 1% having analytics editors on staff (Betts et al. 2017). The survey further revealed that 82% of newsroom positions remained in traditional roles such as reporters, editors, and editorial writers, while journalists generally demonstrated limited proficiency in digital skills. By 2024, however, the International Center for Journalists (ICFJ)[2] predicted substantial transformations in both social and traditional media, asserting that "there will be major changes in social and traditional media, AI disruption will be everywhere, funding for traditional media will evolve"(Newman et al. 2024).

1 OSINT is an abbreviation of Open Source Intelligence - a practice of gathering and analyzing publicly available information to check it for disinformation

2 https://www.icfj.org/news/3-predictions-media-2024, last accessed 9.4.2025

The discourse surrounding artificial intelligence in journalism gained significant momentum in 2023, when the World Press Institute and various partners convened the "New Horizons in Journalism" conference in Sofia. This event facilitated critical exchanges between journalists and media professionals regarding AI-related risks, with particular emphasis on the interplay between journalistic practice and artificial intelligence technologies. While participants acknowledged AI's dual potential in both propagating and detecting misinformation, as well as its applications in investigative journalism and content personalization, they consistently emphasized the enduring importance of human editorial judgment amid challenges posed by algorithmic biases and deepfakes.[3] Notably, journalism, historically characterized by slow adaptation to technological innovations, has begun proactively embracing large language models (LLMs) and generative AI, recognizing their capabilities. The Open Society Foundation's AI in Journalism Challenge (AIJC) exemplifies this shift, providing funding, mentorship, and educational resources to 12 innovative, digital-first newsrooms globally to develop pragmatic AI projects.[4] Preliminary observations from the AIJC indicate that hands-on engagement with AI tools rapidly enhances teams' understanding and enthusiasm, while suggesting that journalistic expertise frequently proves more valuable than technical knowledge in determining project success.

The latest Reuters Institute Digital News Report (Suárez et al. 2024) documents significant shifts in news consumption patterns, with increasing preference for private messaging and video platforms, while maintaining access via social media, search engines, and content aggregators. Only approximately 22% of consumers primarily utilise news websites or applications, representing a notable decline from 32% in 2018. The report identifies several key trends, including the growing influence of content creators, the increasing presence of AI in public discourse, and the introduction of "new layers to news production and distribution" with which journalists and audiences continue to grapple. Particularly concerning is the "proliferation of AI 'slop', low-quality, mass-produced content designed for clicks," occurring alongside democratic backsliding, deteriorating working conditions for journalists, and a rapidly changing digital environment (ibid.). The same report indicates that investigative journalism outlets worldwide are leveraging AI to enhance workflows and expand audience reach through diverse initiatives – from automating social media content creation and summarising articles to analysing complex documents and detecting audio deepfakes – ultimately allowing journalists to focus on in-depth reporting (ibid.). In this context, "platformization becomes increasingly important for how people communicate and access information, including news" (Nielsen and Fletcher 2023). As media organisations increasingly establish agreements with AI companies, it is noteworthy that "most of the media managers didn't expect that money would be shared equally across publishers" (Newman et al. 2024). Trust in traditional media has declined (Newman et al. 2024), particularly in countries experiencing diminished television news consumption and increased social media news usage.

3 https://aej-bulgaria.org/new-horizons-in-journalism-2023-summary/, last accessed 9.4.2025

4 https://www.journalismfestival.com/programme/2024/applying-ai-in-small-newsrooms.-lessons-from-the-ai-in-journalism-challenge, last accessed 9.4.2025

Integrating AI into journalistic practice will catalyse numerous transformations, including automated content analysis, novel journalistic formats, evolving platform-media relationships, implications for privacy and transparency, regulatory developments, ethical considerations, emerging professional roles, and shifts in journalist training (Veleva 2024). Researchers have critically examined issues of transparency, accountability, and bias in AI systems, emphasising the necessity for ethical standards in AI-driven journalism (Broussard et al. 2019),(Opdahl et al. 2023), (Verma 2024). While AI will not supplant human journalists or journalistic intuition, "the new technology brings science to storytelling, helping newsrooms scale their production" (Marconi 2020). Core journalistic competencies – such as editorial judgment, understanding audience needs, identifying and verifying stories, and communicating effectively – will remain essential, requiring new skills in abstract thinking and analytical audience understanding (Veleva 2024). Some researchers have demonstrated that so-called Constitutional AI offers potential solutions to existing challenges in journalism by customizing LLMs to address misinformation and rebuild reader trust (Cheng 2025).

The literature also addresses AI's transformative role in journalism and the emergence of new paradigms such as open data journalism, big data journalism, blockchain journalism, and cloud journalism (Hassan and Albayari 2022). Scholars have investigated public perception of AI-generated news content and humans' ability to distinguish between articles authored by humans versus those created by artificial intelligence (Moravec et al. 2024). Research on AI's capabilities in composing various journalistic texts – both informative and opinion-based – reveals that common applications include automatic content generation, data analysis, documentation, and text translation, while usage decreases for creative tasks such as headline writing or advanced editorial functions (Fernández-Barrero, López-Redondo, and Aramburú-Moncada 2024). The term "algorithmic journalism" has emerged to describe technological transformations in the field, defined as "a process of using software or algorithms to automatically generate news stories without human intervention" (Graefe 2016). Some scholars conceptualise algorithmic journalism more broadly, encompassing automated content production, data mining, news dissemination, and content optimisation (Kotenidis and Veglis 2021).

Although not the primary focus of this research, natural language processing for fact-checking represents a critical area for future investigation. Such research should "include collaboration with fact-checkers, as well as incorporation of human-centred design practices in model development, to further guide technology development for human use and practical adoption"(Das et al. 2023).

As AI solutions continue to demonstrate potential for transforming newsrooms by enhancing efficiency, accuracy, and accessibility in news production, scholars have raised important questions regarding tool ownership and transparency. Research indicates that "the lack of transparency is a significant concern, particularly with regard to the transparency of AI tools utilized for fact-checking information in journalism: of the 100 AI tools identified, 23 included AI fact-checking services, and of these 23, only five (21%) could be classified as adequately transparent. 13 of them, or over 56% of the total, are considered not transparent" (Martin 2024). The advantages of AI could lead to slow journalism, including generating and enhancing content, reducing workloads, and consequently giving journalists more time for creative tasks (Albizu-Rivas, Parratt-Fernández, and Mera-Fernández 2024).

3. Three Types of Possible AI Tools and Their Uses in Bulgarian

3.1 General Observation for Bulgaria

Shi and Sun summarise three main applications of AI in journalism: information gathering; content production; customization and dissemination (Shi and Sun 2024). First, monitoring and gathering of news and the analysis of data can forecast trends and prepare journalists for emerging events (like Reuters News Tracer filtering Twitter for breaking news), addressing journalists' time and space constraints by providing real-time information monitoring, and being trained by journalists to align with editorial standards (ibid.). Second, AI can assists journalists in content creation by generating text that matches news organizations' style and tone, managing tasks as facilitating transcription, translation, and improving readability, producing multilingual content, expanding audience reach – demonstrated by Le Monde's AI-assisted translation of approximately 30 stories daily for its English edition, and NewsGPT (launched March 2023), which analyzes data from multiple sources to create reports using AI anchors. AI supports content production in two key ways: first, by creating summaries and headlines and converting between text, images, audio, and video formats; second, by diversifying writing styles—adapting stories for different publications and audiences, as shown Claude's ability to narrate news in styles mimicking various major publications (ibid.). Third, customisation and dissemination, where the traditional news aims at a specific audience with previously analysed preferences (ibid.).

When conducting a comparative analysis of technological trends in Bulgarian media landscapes versus international standards, several significant disparities emerge. The Bulgarian media ecosystem demonstrates notable deficiencies in technology across multiple dimensions of journalistic practice. Developing a mini-framework for evaluating AI tools for journalism, we analyse language support, transparency and usability of the tools.

The Bulgarian media landscape is characterised by a persistent lack of transparency in both ownership structures and editorial practices, despite the existence of legal frameworks intended to promote openness (Bleyer-Simon et al. 2024). Within the journalistic community itself, there is a notable division regarding professional standards and responses to disinformation, particularly as digital transformations and debates over misinformation dominate the public discourse. Actors from across the political spectrum actively engage in these debates, each asserting the legitimacy of their own narratives. Many websites, which name themselves as news websites, use copy-paste journalism (see below the journalistic investigation about the so-called mushroom news websites).

Every journalist, who attends an event, must create text. Again, depending on the subject matter, journalists use various open public data for verification, although in Bulgaria, so-called data journalism is not well developed. When dealing with non-international news, journalists work in direct connection with the main "subjects" of their news texts. They can directly verify certain information from the primary source and conduct direct journalistic investigations. Despite the existence of various tools advertised as facilitating journalists' work, journalists in Bulgaria still rely on their own capabilities and journalistic intuition. There are several reasons for that – one is the lack of understanding in many newsrooms about the need to introduce technology and the lack of sufficient information about new

technologies. Journalists are largely self-taught in terms of technological advances. On the other hand, each journalist develops their verification channels and personal methodology for checking specific cases. In this sense, every journalist, in the classical meaning of the word journalist, possesses both a solid background in a given field as well as contacts and their personal strategy. In Bulgaria, in the recent past, there was an entire specific team of librarians in every newspaper editorial office, called Documentation, where journalists were assisted, verifying data and conducting comprehensive research about specific historical events and individuals. These departments have now been closed, and for data verification, journalists rely, as already mentioned, on their own resources – knowledge, experience, and instinct.

3.2 Monitoring and Gathering

Primary media organizations in Bulgaria predominantly lack automated monitoring systems, relying instead on manual monitoring of the news flow. The technological tools widely utilized for English-language content processing have minimal Bulgarian-language equivalents, creating a substantial capability gap for journalists working in this linguistic context.

Creating a journalistic article depends on the field to which it is dedicated. As is well known, journalists typically specialize in different fields and develop in-depth knowledge of their subject matter. However, with the emergence of large online platforms and social media, journalism has become increasingly dynamic, requiring much faster reactions.

We will skip the deep analysis of the various most popular browsers on the internet, since the results for the Bulgarian language in terms of news are close, and each journalist chooses which one to work with, as may be the editorial policy. In this regard, we could expect that some AI-powered tools could be useful, which could synthesise information from the web and providing concise, well-sourced responses in the Bulgarian language. Some of these tools are checked in Bulgarian and provide efficient research, fact-checking, and quick access to background information, as well as summarising articles, generating topic overviews, and suggesting related content streamlines (ex. Perplexity).[5]

In Bulgaria, journalists monitor their own information – they follow the agenda of the parliament, the council of ministers, the work of their colleagues. Each editorial office has a subscription to certain agencies and channels, as well as to news exchanges. A huge number of Bulgarian technology companies offer media monitoring, press clipping, internet clipping, and media analysis, but these possibilities are used by specific actors, not by journalists daily, and they monitor the media flow in an "analog" way – checking other news websites, monitoring national televisions (here we are referring to transmedia journalism, in which news from one type of media is transferred to another type of media), monitoring the Bulgarian Telegraph Agency (which often does not get ahead of the news flow), and certain social media groups and accounts.

5 https://www.perplexity.ai/

However, there are some free options for monitoring information in Bulgarian, although these technological tools are much slower than journalists manually reviewing websites, and they are news aggregators, which collect and republish online already published news from various sources. One of them is Google's news aggregator[6], created back in 2002, which also works with Bulgarian. The user can explore the aggregator by their Google account, and as stated in the announcement of company, the: news articles are ranked based on their quality, originality of content, freshness of content, and where permitted based on settings and previous activity and purchases within Google News, and activity in other Google products. Google may have a license agreement with some of these publishers, but it has no impact on the ranking of results. Even having that disclaimer, it is not transparent how the algorithm works. That lack of transparency of the algorithm appears in all news aggregators.

The advantage of this aggregator is that the user can choose the sources to follow. Thus, having enough experience with reliable and unreliable sources, journalists choose which source to follow. However, this can be a problem when the journalist wants to catch disinformation and oppose it before it even takes on gigantic proportions. Another advantage of this aggregator is that the topics that the journalist follows can be indicated, and at the same time, get an overview of the entire flow, following Top Stories. A third advantage is that the aggregator does not offer summarised news, but original articles and the editor can choose where to focus on. The aggregator has a search field, where the journalist could search by keywords, Boolean search, and period. However, unlike the same aggregator in English, where there is also a fact-checking section, the Bulgarian tool lacks this functionality.

Over the years, there have been various aggregators only in Bulgarian, for which it was not clear who created them, and most of them are no longer available. The Topnovini[7] aggregator is of this type. It collects various news, distributed by topics. Through the settings, the user can choose from 17 websites that work in Bulgarian. There is no field for searching by keywords, nor for a certain period. There is a ranking of the most-read news in the different categories. These functionalities are not available with all browsers. The aggregator is updated every 10 minutes. Despite the transparency of the way that news was collected, the ownership of the website is not clear, and a check at https://hostingchecker.com/ gives results that the site is hosted by Hetzner Online GmbH, Helsinki, Finland.

Another similar Bulgarian news aggregator is radar.bg. [8] It offers 20 different subtopics from which the user can choose which ones to follow. It seems that the news stream is updated every hour, which is not particularly convenient for the dynamic news environment. The channel allows keyword search. The aggregator provides a link to the articles, as well as the first 500 characters of each article, without summarising them. There is a lack of transparency about what news websites it is fed to, as well as who is behind it. A check at https://hostingchecker.com/ shows that it is hosted by: MAIL.BG Ssc, Sofia, Bulgaria.

6 https://news.google.com/home?, last accessed 8.3.2025
7 https://www.topnovini.com/, last accessed 8.3.2025
8 https://radar.bg/news

The tool NewsGPT[9] is not relevant for the Bulgarian environment. The news related to Bulgaria is from months ago and cannot work as a tool for monitoring and gathering news. The slogan "Tomorrow news today", implementing the AI predictions of news is not working in Bulgarian and it could be used only for amusement, based on the logo *News by AI. Share the unhuman truth.* The Google Trends tool also works for the Bulgarian language. It is convenient to get a very general picture of individual searches in Google, providing statistical data on searches both on the territory of Bulgaria and by regions of Bulgaria. As the company says, it shows what's trending across Google Search, Google News and YouTube.

Many of the world's news agencies have tools for personal notification of breaking news, but the content is in English and is of interest mainly to international news editors.

3.3 Content Creation

The identified technological deficiencies have substantive implications for journalistic practice across multiple dimensions. Limited technological capacity compromises information verification processes, particularly in environments characterized by high information velocity and sophisticated misinformation. Content creation and content processing are the most interesting from a linguistic point of view. Without underestimating the data collection itself, which is sometimes a real journalistic investigation, the creation of content itself is perceived as a creative process.

There have been multiple developments in speech and language technology, and Bulgarian is part of some multilingual systems for machine translation, speech analysis and recognition (Koeva 2023). A Bulgarian General Language Understanding Evaluation Benchmark - bgGLUE, was also created, for evaluating language models on natural language understanding tasks in Bulgarian, targeting a variety of NLP problems (e.g., natural language inference, fact-checking, named entity recognition, sentiment analysis, question answering, etc.) and machine learning tasks (sequence labeling, document-level classification, and regression) (Hardalov et al. 2023). In recent years, many NLP scholars have been working specifically on the topic of disinformation in Bulgarian (Hardalov, Koychev, and Nakov 2016), (Koeva 2021), (Nakov et al. 2021), (Osenova and Simov 2024), (Temnikova et al. 2023), (Margova 2023). Disinformation is not the main focus here, however, when talking about content, fact-checking in journalistic content is a basic norm. In the autumn of 2024, the Association of European Journalists Bulgaria (AEJ) and the lisenced by European Fact-Checking Standards Network (EFSCN)[10] fact-checking organisation FactCheck.bg are launching a new partnership to combat election disinformation as a part of Google News Initiative, with a main goal of fighting misinformation and monitoring electoral fraud.[11] In 2018, Google launched its News Initiative to scale the work with journalists, publishers, and

9 https://newsgpt.ai/
10 https://efcsn.com/
11 https://aej-bulgaria.org/en/google-news-initiative-aej-bulgaria-launch-a-new-partnership-against-election-fraud-and-disinformation/, last accessed 3.8.2025

industry leaders to help build a resilient future for news. Thus, out of the entire package of different Google News Initiative capabilities, the fact-checking part has a Bulgarian version and is powered by Bulgarian verified facts.

In the context of fact-checking and relevant debunkings across languages, the Database of Known Fakes (DBKF)[12] allows users to check whether a claim, image or video has already been verified by trusted sources, when and how, using AI-powered technologies that go way beyond a keyword search. This could be particularly useful for fact-checkers and journalists as part of their verification workflow. The Bulgarian part consists of more than 2000 debunked articles. The database is organized by claims and articles.

Another similar database at the European level, containing identical articles in Bulgarian, like that of DBKF, is the Truly media[13] platform, which is, however, paid. In Bulgarian, a number of media outlets are engaged in fact-checking, with the licensed ones being under the umbrella of the BROD project – Bulgarian National Television Bnt[14] and AFP[15], and Factcheck.bg[16]. These details are important because, as noted above, too often, the disseminators of disinformation accuse others of such actions.

Another important tool that can be used in fact-checking is the InVid[17] plugin for Google's Chrome browser. This toolkit is a "Swiss army knife" helping journalists save time and be more efficient in their fact-checking and debunking tasks on social networks especially when verifying videos and images. It is important that the tool also works in Bulgarian and it is able to analyse the intensity of emotions in a given text. Nowadays, new automated fact-checker tools[18] have appeared and some of them performed quite well in Bulgarian, but it is a fast-developing field and many improvements are needed. Special evaluator system is created (Wang et al. 2024).

Unfortunately, the other possibilities available to journalists within the framework of the Google News Initiative are not accessible in the Bulgarian language. However, it should be noted that the programme offers a number of courses for journalists related to the use of AI, as well as machine learning, LLMs, and offers a Pinpoint workspace in which files can be easily transcribed, but these options are not working in Bulgarian.

Microsoft also has experience in creating healthy news ecosystems, explaining how technology played a role in the disruption of news, but can also be an important part of the rebuilding effort. Thus, Microsoft, and its Democracy Forward programme, provide a host of tools[19] and services to help journalists to rebuild capacity, restore trust, and reduce risk. As in the case with Google, these tools are not available in Bulgarian.

12 Developed in the frame of the European projects WeVerify and BROD project by Ontotext https://brodhub.eu/en/fact-checking/ontotext-dbkf/
13 https://www.truly.media/
14 https://bntnews.bg/bnt-provereno-108929tag.html
15 https://brodhub.eu/bg
16 https://factcheck.bg/
17 https://www.invid-project.eu/tools-and-services/invid-verification-plugin/
18 https://www.longshot.ai/fact-check-free, https://www.factiverse.ai/, https://originality.ai
19 https://www.microsoft.com/en-us/corporate-responsibility/journalism-hub

A huge number of initiatives like Partnership on AI (PAI)[20] for example tries to connect academic with civil society, industry, and media organizations to create sustainable development for journalists in the age of AI, but they give general recommendations, and not the exact tools, especially in Bulgarian.

Several specific tools work for the Bulgarian language and are somehow helpful for journalists. Various spell-checking programmes are now taken for granted, even for the Bulgarian language. However, one big issue remains in word processing – the so-called transcribing of sound files, as writing journalists continue to work with them. Unfortunately, most free tools do not support the work of journalists, as their use is limited and they end up having to manually download the content of the sound files. The only possible option is the Turboscribe[21] tool, which has a free version and allows processing up to 3 audio and video files in 24 hours, where a GPU-powered transcription engine converts audio and video to text, to be exported as DOCX, PDF, TXT, captions, and subtitles (SRT, VTT). However, the result is full of errors and must be manually checked by the journalist in questionable places, and editing takes time.

Another major need is related to translations, especially when it comes to covering international news. Journalists use various free online translators, which in recent years have been getting better and better at handling the Bulgarian language. Naturally, Google Translate[22] and Deepl[23] are mainly used. They are convenient for short texts. Deepl handles the Bulgarian language especially well, as it also allows for additional editing, by containing different suggestions. Even not very popular, Gourmet[24] is other tool which can be used for translation from and to Bulgarian, with language pairs or models for it. Although international news and the translation of news articles are not a priority, translator tools provide great opportunities when it comes to languages unfamiliar to the journalist and for conducting journalistic investigations. However, the main shortage of journalists during investigations remains time. Translation tools could also be used for creating content in English to make news websites more visible outside Bulgaria.

The main game changer nowadays – large language models (LLMs) and AI tools, such as bots, could automate and enhance news production. To date, no research has been conducted on whether and to what extent LLMs and chatbots have been used for content creation in Bulgarian media, and for what kind of tasks. Only sporadic freelancers share their news-generating experiences and one media shares such a rubric.[25] Tools like BGGPT[26], ChatGPT[27], Claude[28], Deepseek[29], Gemini[30] could assist content creation, research, data analysis, summarising translations, fact-checking, and editorial decision-making. These

20 https://partnershiponai.org/resources/
21 https://turboscribe.ai/
22 https://translate.google.com/?hl=bg
23 https://www.deepl.com/en/translator
24 https://gourmet-project.eu/project-output/gourmet-translate-tool/
25 https://karamanev.me/ and https://clubz.bg/151538
26 https://bggpt.ai
27 https://openai.com/index/chatgpt/
28 https://claude.ai/
29 https://chat.deepseek.com/
30 https://gemini.google.com

agents could also personalise news delivery by analysing reader preferences, suggesting related topics. All the outputs require human oversight to ensure accuracy and uphold journalistic standards. Claude has a similar application as ChatGPT, but is aligned with ethics, minimising harmful bias, while Deepseek is blamed for propaganda dependencies.[31] Generally speaking, drafting articles, summarising reports, and translating are among the main possibilities that AI tools offer.

3.4 Dissemination

Bulgarian media demonstrates limited implementation of artificial intelligence systems for audience segmentation and content personalisation. While media organisations in many countries employ sophisticated algorithms to tailor content delivery according to user preferences and behavioural patterns, Bulgarian outlets have not substantively integrated these capabilities into their operational frameworks.

Editors of news websites are usually also involved in the distribution of content on various social platforms. All major media outlets have their accounts on the various social networks popular in our country. Publications are made manually, although there are also many content optimisation experts who can assist journalists. Again, it comes down to the issue of enriching experience and knowledge in harmony with new technologies.

In Bulgaria, dissemination is a priority for malicious actors in the field of journalism because it's directly linked to monetisation. Some individuals without journalistic education can still influence social life within the country. For example, the Bulgarian journalist Georgi Angelov[32], made an investigative reportage, in which he became part of the so-called disinformation machine - part of "mushroom cites". Angelov managed to become involved in a scheme that disseminated disinformation, as well as specific articles about certain politicians designed to shape public opinion in Bulgaria. He manages to expose the system of disinformation and to show how it works. The creation of "mushroom sites" that operate in favour of foreign political interests and threaten a country's national security should not remain solely within serious online media outlets. The reason for highlighting this case is the possibility for any individual to become a journalist by joining such a mushroom site machine and monetising their work directly—that is, although it constitutes political propaganda, disinformation, and actions that are dishonest from society's perspective, this represents a scheme in which technologies play a role and, with human assistance, create false misleading content that contradicts the law. This raises the question of ethics in using various tools, which is essential and must be considered in journalism.

31 https://www.politico.eu/article/we-asked-deepseek-about-geopolitics-chinese-government-propaganda-artificial-intelligence/

32 https://www.svobodnaevropa.bg/a/saytove-gabi–mashina–dezinformatsiya-rusiya/32889599.html, last accessed 1.4.2025

3.5 Analysis

The existing tools for monitoring and collecting information in Bulgarian on the Bulgarian Internet can provide a fairly general picture of what is happening in the country, but they would not adequately serve journalists, who are certainly ahead of the events being covered. In addition to the delay, which is critically important in the profession, there is also a lack of transparency of the algorithms behind the various aggregators, as well as their owners. This can be a problem in times of disinformation and impaired information integrity. At the same time, as already mentioned, many Bulgarian technology companies are engaged in media clipping and monitoring, and it is possible to assist journalists with free, simplified versions of their tools in the name of a better journalistic environment that is more stable against disinformation attacks.

Bulgarian transcription tools are still a limitation for journalists. Large language models are limited in time (the data on which they are trained is limited to a certain date) and cannot be used to check today's news. However, they can provide ideas. A main problem when working with AI tools is that the content they generate must be checked, and in the hectic journalistic everyday life, time is one of the main resources. Although our emphasis is not on fact-checking, this is a field that is developing, and Bulgarian journalists can use it. The same applies to new search engines. It is important to note that many of the opportunities that exist in the field of journalist education are not offered in Bulgarian, and are not particularly well-known among media managers. Training journalists is an opportunity to improve the media environment as a whole. Despite the limited time of journalists, newsrooms can invest in their knowledge and upgrade their qualifications regarding AI, which will contribute to the overall efforts against disinformation. Since the so-called data journalism is not developed in our country, the opportunities provided by chatbots and AI tools can contribute to the development of this data journalism, as they can easily work with large data sets. Many automated fact-checker tools have appeared and need to be deeply analysed for Bulgarian.

Bulgarian media exhibits a notable lag in adopting AI for audience segmentation and content personalisation, contrasting with global trends. This limited integration hinders audience engagement, advertising effectiveness, and the ability to understand evolving user needs. The manual distribution of content on social platforms, despite the availability of optimization expertise, further underscores this technological gap.

Compounding this challenge is the exploitation of these very technological limitations by malicious actors. The case of "mushroom sites" demonstrates how the ease of online content creation and dissemination, coupled with a focus on monetisation, enables the spread of disinformation and politically motivated propaganda. This not only undermines journalistic integrity but also poses a threat to national security. The ethical considerations surrounding the use of technology in journalism, particularly in creating and amplifying misleading content, become crucial. It is paramount that trustworthy media outlets, committed to genuine journalism, achieve greater reach to effectively counter the noise generated by disinformation sources. In an era saturated with information, discerning credible news from falsehoods is increasingly challenging for the public. Amplifying the voices of reputable

journalists and news organisations is crucial for fostering an informed citizenry capable of making sound decisions.

These trusted sources adhere to journalistic ethics, prioritise factual accuracy, and provide well-researched analysis, offering a clear contrast to the often emotional and fabricated content propagated by disinformation networks. By strategically leveraging digital platforms and innovative engagement strategies, credible media can penetrate echo chambers and connect with wider audiences. Investing in media literacy initiatives can further empower individuals to critically evaluate information and identify manipulation tactics. Ultimately, ensuring the prominence of authentic journalism is vital for safeguarding democratic values and fostering a society grounded in truth and informed discourse, effectively drowning out the cacophony of misinformation.

4. Ethical Questions

Ethics in journalism is a main question. The integration of artificial intelligence tools in Bulgarian journalism raises significant ethical questions that require careful consideration. While technological advancements offer numerous benefits for news production and distribution, they simultaneously introduce complex ethical challenges that must be addressed through robust frameworks and standards.

4.1 Transparency and Accountability

Without diving deeper, here we will mention the need to develop ethical standards for the use of AI. European Union made a big step in that direction, developing the Digital Service Act, where the Code of Practice on Disinformation is recognised as a robust set of commitments for Very Large Online Platforms (VLOPs) and Very Large Online Search Engines (VLOSEs) to constitute strong mitigation measures against online disinformation, as demonetisation: cutting financial incentives for purveyors of disinformation; transparency of political advertising: more efficient labeling for users to recognize political advertising; ensuring the integrity of services: reducing fake accounts, bot-driven amplification, malicious deep fakes, and other manipulative behavior used to spread disinformation; and empowering users, researchers, and the fact-checking community with better tools for users to identify disinformation, wider access to data, and fact-checking coverage across the EU. This framework encompasses crucial measures such as demonetisation of disinformation, transparency in political advertising, service integrity maintenance, and empowerment of users, researchers, and fact-checkers. However, these broad European standards require adaptation and specific implementation for the Bulgarian media landscape, taking into account local journalistic practices and challenges.

4.2 Attribution and Authorship

The emergence of AI-generated or AI-assisted content creates fundamental questions about proper attribution and intellectual responsibility. When Bulgarian journalists utilise large

language models or other AI tools to produce content, how should readers be informed about AI involvement in content creation? This issue becomes particularly salient in the context of the case study analysis conducted in this research, examining articles where AI is explicitly identified as a co-author of journalistic content.

Without clear standards for attributing AI contributions, it potentially creates confusion among readers, searching for original and reliable information and may erode public trust in media institutions already struggling with credibility challenges in Bulgaria's complex information environment.

4.3 AI Slop and Journalistic Standards

AI tools could be efficient but may compromise journalistic quality if implemented without appropriate oversight mechanisms. In the resource-constrained environment of Bulgarian newsrooms, where time pressures are significant and technological expertise is limited, AI tools might be deployed without sufficient quality control processes. The "AI slop" phenomenon identified in the Reuters Institute Digital News Report (Suárez et al. 2024) – low-quality, mass-produced content designed primarily for click generation – represents a concerning potential outcome when AI implementation prioritises quantity over quality. The best practice is to maintain journalistic standards, creating clear editorial policies about AI usage, including protocols for human oversight, fact verification, and editorial judgment.

4.4 Data Privacy and Consent

The work with a large amount of data is not eas, and Bulgarian journalists generally do it manually. When journalists utilise AI tools for content analysis, audience segmentation, or personalisation, they inevitably engage with significant amounts of user data, the questions about privacy protection raised, particularly in the context of Bulgaria's implementation of the General Data Protection Regulation (GDPR). Media organisations must ensure that data collection and processing practices comply with legal requirements, respecting audience privacy. Bulgarian journalists must navigate these boundaries thoughtfully, considering both legal compliance and ethical responsibility.

4.5 Misinformation and AI Detection

As AI-generated content becomes increasingly sophisticated, distinguishing between human-authored and AI-generated texts presents growing challenges for Bulgarian journalists and media consumers. The potential for malicious actors to employ AI tools for creating and disseminating misinformation, as illustrated in the "mushroom sites" investigation discussed earlier, highlights the dual nature of AI technologies in the information ecosystem.

Developing effective strategies for detecting AI-generated content and countering misinformation requires specialised knowledge and tools that many Bulgarian media organisations currently lack. This technological gap creates vulnerabilities in the information environment that could be exploited to undermine public discourse and democratic processes.

5. Case Study: *AI Reporting Live from the Scene*

In July 2024, the Bulgarian online media outlet ClubZ announced that it was starting a project in which editors would test and develop ethical use of artificial intelligence in journalism under the name "AI reporting live from the scene". [33] From July 2024 to March 2025, 172 articles were published in this section, co-authored by artificial intelligence. Thus, an average of 20 articles per month are created by AI. Of all the titles, only 16 are related to Bulgaria, all the rest concern foreign policy events or are related to artificial intelligence and its implementations. Those related to Bulgaria are distributed by topic, respectively: most are related to the Ministry of Defence, a few political articles related to the formation of a government, and a few for consumers. In all articles, the AI is indicated as the author, without having a name of journalists, and at the end of the text there is a disclaimer reflecting the co-authorship of artificial intelligence: „This material was written with the help of artificial intelligence under the control and editing of at least two journalists from Club Z. The material is part of the project "AI reporting live from the scene". The rest of the articles demonstrate a thematic preponderance of AI, with Elon Musk's name appearing across discussions related to technological advancements, legal frameworks, and innovative methodologies.

Geopolitical tensions and cybersecurity concerns, particularly those involving China, Russia, and the United States, constitute another prominent trend. Furthermore, a substantial portion of the articles addresses the legal and ethical implications of AI's societal impact. While less frequent, articles dedicated to pure technological innovation, without political or legal framing, remain a smaller but still visible topic.

In addition to manual discourse analysis, we used ChatGPT for a quick analysis of the topics in the articles. The check shows that the disclaimer containing AI reporting live from the scene is perceived as a permanent topic for artificial intelligence, and according to ChatGPT, all articles contain this topic. However, this is not the case, because the texts related to Bulgaria are not aimed at artificial intelligence, but at completely different topics. The tool's error proves that although very good for big data analysis, tools like Chat GPT should be used carefully and checked. Chat GPT does not recognise topics related to Bulgaria as a different topic at all. According to the tool, the next main topic is geopolitics and security, followed by the work of technology giants, the actions of Elon Musk, Social & Legal Issues, and Innovation & Future Tech. The experiment they are doing at Club Z is commendable for several reasons – the editorial team dares to use AI, tries to introduce ethical norms into this use, and publicises it. Proportionally, based on this experiment, it becomes clear that the use of AI is currently more applicable to international news and topics related to AI. One possible explanation is that Bulgarian topics require editorial intervention and scrutiny, especially when it comes to politics, defence, and consumer interests. Nevertheless, the experiment is worth it. So far, only freelancers have been doing such experiments with text generation. As a potential recommendation, we would suggest greater transparency regarding the type of AI employed and the specifics of its application. Clearly outlining how AI tools

33 https://clubz.bg/151538

are integrated into journalistic processes can foster public trust and understanding. This openness can help distinguish ethical AI usage from potentially manipulative applications, contributing to a more informed media landscape.

6. Conclusions and Recommendations

This research underscores that while significant technological gaps exist in the Bulgarian journalistic landscape, targeted investment, training, and ethical framework development could substantially enhance the sector's capacity for rigorous reporting and strengthen its institutional sustainability in an increasingly complex information environment.

Bulgarian media organisations demonstrate notable deficiencies in technological adoption across multiple dimensions of journalistic practice, particularly in comparison to international counterparts. Primary media organisations in Bulgaria predominantly rely on manual monitoring rather than automated systems. Available news aggregators (Google News, Topnovini, Radar.bg) offer limited functionality, lack transparency in their algorithms, and update too infrequently for the dynamic news environment. Perplexity AI represents a promising but underutilised resource for Bulgarian journalists. Bulgarian journalism faces significant challenges in content processing, with limited AI-powered tools specifically developed for the language. While spell-checking programmes exist, transcription tools for audio files remain inadequate. Translation tools like DeepL show promise but have limitations for journalistic investigations. Large language models (BGGPT, ChatGPT, Claude, Deepseek, Gemini) offer potential assistance with content creation but require further evaluation and integration into newsroom workflows.

Several fact-checking initiatives operate in Bulgarian, including those under the BROD project (Bulgarian National Television, AFP) and Factcheck.bg. Tools like InVid and databases such as the Database of Known Fakes provide some support for verification processes, though automated fact-checkers for Bulgarian require further development. Bulgarian media demonstrates limited implementation of artificial intelligence systems for audience segmentation and content personalisation. Trustworthy journalism must reach wider audiences to drown out disinformation. In today's information overload, distinguishing fact from fiction is a growing challenge. Amplifying reputable news sources, committed to ethics and accuracy, is vital for an informed public. Strategic use of digital platforms and audience engagement can help credible media break through echo chambers.

While the European Union's Digital Service Act (DSA) provides a broad framework for addressing disinformation, more specific ethical standards for AI use in Bulgarian journalism are needed to ensure responsible implementation of the DSA. Media organisations should prioritise investment in AI tools specifically developed or adapted for Bulgarian language processing, particularly in areas of transcription, automated monitoring, and content analysis. New comprehensive training initiatives for journalists on effectively utilising AI tools in their workflow, with special attention to language-specific capabilities and limitations, are needed. AI tools could be used for trustworthy data journalism practices in Bulgaria, an underdeveloped area with significant potential for enhancing reporting quality and depth.

It is necessary to create specific standards for ethical AI use in Bulgarian journalism, addressing the unique challenges of the local media landscape while aligning with broader European frameworks. Media organisations should collaborate with technology companies to develop more effective Bulgarian-language AI tools tailored to journalistic needs. Addressing ethical challenges requires an approach involving media organisations, technology developers, regulatory bodies, academic institutions, and civil society organisations.

Specific actions include: developing ethical guidelines for AI implementation in Bulgarian journalism, adapted to local conditions; establishing transparency requirements for AI tools used in newsrooms; creating standard practices for AI-assisted or AI-generated content to ensure audience understanding of content provenance; implementing robust quality control mechanisms to prevent the proliferation of AI slop; providing specialised ethical training for journalists and editors on responsible AI usage; fostering cross-sector collaboration between media and technology companies to develop AI tools that support journalistic values. These standards could contribute to the ethical integration of AI in journalism, supporting democratic discourse and enhancing media credibility.

Implementing the recommendations above, Bulgarian journalism would be a tech-empowered, ethically grounded media ecosystem, where AI tools tailored to the Bulgarian language would support the journalistic process. Receiving training, helping human editorial judgment with the precision of machine assistance, and approving ethical standards would rebuild public trust in the media, and data journalism would make disinformation fade, leading to information integrity.

Acknowledgments

The author thanks the reviewers of the journal for their remarks. The author is also grateful to his colleagues Irina Temnikova and Venelin Kovatchev. This research has been supported by funding from the GATE project funded by the Horizon 2020 WIDESPREAD-2018-2020 TEAMING Phase 2 Programme under grant agreement no. 857155, the Programme "Research, Innovation and Digitalisation for Smart Transformation" 2021-2027 (PRIDST) under grant agreement no. BG16RFPR002-1.014-0010-C01, and the project BROD (Bulgarian-Romanian Observatory of Digital Media) funded by the Digital Europe Programme of the European Union under contract number 101083730.

References

Albizu-Rivas, I., S. Parratt-Fernández, and M. Mera-Fernández. 2024. Artificial intelligence in slow journalism: Journalists' uses, perceptions, and attitudes. *Journalism and Media*, 5(4):1836–1850.

Betts, T., D. Bowers, J. Cohen, L. Evans, C. Gicheru, M. Jenkins, R. Khursheed, A. Mhanna, H. Owusu, M. Peralta, S. Vucinic, and M. Zayani. 2017. Icfj survey: The state of technology in global newsrooms. Technical report, ICFJ Survey.

Bleyer-Simon, K., E. Brogi, R. Carlini, D. Da Costa Leite Borges, J. Kermer, I. Nenadic, M. Palmer, P. Parcu, R. Reviglio, M. Trevisan, S. Verza, and M. Žuffová. 2024. Monitoring media pluralism in the digital era: Application of the media pluralism monitor in the european member states and in candidate countries in 2023. Technical report, Centre for Media Pluralism and Media Freedom, Robert Schuman Centre.

Broussard, M., N. Diakopoulos, A. L. Guzman, R. Abebe, M. Dupagne, and C. H. Chuan. 2019. Artificial intelligence and journalism. *Journalism & Mass Communication Quarterly*, 96(3):673–695.

Cheng, S. 2025. When journalism meets AI: Risk or opportunity? *Digital Government: Research and Practice*, 6(1):1–12.

Das, Anubrata, Houjiang Liu, Venelin Kovatchev, and Matthew Lease. 2023. The state of human-centered nlp technology for fact-checking. *Information Processing & Management*, 60(2).

Fernández-Barrero, M. Á., I. López-Redondo, and L. G. Aramburú-Moncada. 2024. Possibilities and challenges of artificial intelligence in the teaching and learning process of journalism writing: The experience in spanish universities. *Communication & Society*, pages 241–256.

Graefe, A. 2016. Guide to automated journalism. Technical report, Tow Center for Digital Journalism.

Hardalov, M., I. Koychev, and P. Nakov. 2016. In search of credible news. In *Artificial Intelligence: Methodology, Systems, and Applications: 17th International Conference, AIMSA 2016, Varna, Bulgaria, September 7–10, 2016, Proceedings*, volume 17, pages 172–180, Springer International Publishing.

Hardalov, Momchil, Pepa Atanasova, Todor Mihaylov, Galia Angelova, Kiril Simov, Petya Osenova, Veselin Stoyanov, Ivan Koychev, Preslav Nakov, and Dragomir Radev. 2023. bgGLUE: A Bulgarian general language understanding evaluation benchmark. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 8733–8759, Association for Computational Linguistics, Toronto, Canada.

Hassan, A. and A. Albayari. 2022. The usage of artificial intelligence in journalism. In *Future of Organizations and Work after the 4th Industrial Revolution: The Role of Artificial Intelligence, Big Data, Automation, and Robotics*. Springer International Publishing, Cham, pages 175–197.

Koeva, S. 2021. The Bulgarian WordNet: Structure and specific features. *Bulgarian Academy of Sciences*, 8(1):47–70.

Koeva, Svetla. 2023. Language Report Bulgarian. In *European Language Equality: A Strategic Agenda for Digital Language Equality*. Springer, pages 103–106.

Kotenidis, E. and A. Veglis. 2021. Algorithmic journalism—current applications and future perspectives. *Journalism and Media*, 2(2):244–257.

Marconi, F. 2020. *Newsmakers: Artificial intelligence and the future of journalism*. Columbia University Press.

Margova, R. 2023. *Linguistic markers of Fake News*. Ph.D. thesis, Sofia University St. Kliment Ohridski. https://bpos.bg/publication/57700.

Martin, S. 2024. Control beyond code: Ownership and financial insights of AI used in journalism. Technical report, Media and Journalism Research Center (MJRC), Tallinn/London/Santiago de Compostela.

Moravec, V., N. Hynek, M. Skare, B. Gavurova, and M. Kubak. 2024. Human or machine? The perception of artificial intelligence in journalism, its socio-economic conditions, and technological developments toward the digital future. *Technological Forecasting and Social Change*, 200:123162.

Nakov, P., D. Corney, M. Hasanain, F. Alam, T. Elsayed, A. Barrón-Cedeño, and G. D. S. Martino. 2021. Automated fact-checking for assisting human fact-checkers. https://arxiv.org/abs/2103.07769. ArXiv preprint arXiv:2103.07769.

Newman, N., R. Fletcher, C. T. Robertson, A. Ross Arguedas, and R. K. Nielsen. 2024. Reuters Institute Digital News Report 2024. Technical report, Reuters Institute for the Study of Journalism.

Nielsen, R. K. and R. Fletcher. 2023. Comparing the platformization of news media systems: A cross-country analysis. *European Journal of Communication*, 38(5):484–499.

Opdahl, A. L., B. Tessem, D. T. Dang-Nguyen, E. Motta, V. Setty, E. Throndsen, and C. Trattner. 2023. Trustworthy journalism through AI. *Data & Knowledge Engineering*, 146:102182.

Osenova, P. and K. Simov. 2024. Representation of multiword expressions in the Bulgarian integrated lexicon for language technology. In *Multiword Expressions in Lexical Resources: Linguistic, Lexicographic, and Computational Perspectives*. pages 117–146.

Shi, Y. and L. Sun. 2024. How generative AI is transforming journalism: Development, application and ethics. *Journalism and Media*, 5(2):582–594.

Suárez, E., M. Adami, M. Leake, and G. Kahn. 2024. How 2024 shaped journalism: Insights from the Reuters Institute's work. Technical report, Reuters Institute for the Study of Journalism. https://reutersinstitute.politics.ox.ac.uk/news/how-2024-shaped-journalism-insights-reuters-institutes-work, last accessed 8.4.2025.

Temnikova, I., I. Marinova, S. Gargova, R. Margova, and I. Koychev. 2023. Looking for traces of textual deepfakes in Bulgarian on social media. In *Proceedings of the 14th International Conference on Recent*

Advances in Natural Language Processing, pages 1151–1161.

Veleva, M. 2024. Навлизането на изкуствения интелект в медиите – предимства и рискове [The introduction of artificial intelligence in the media – advantages and risks]. *Медии и обществени комуникации [Media and Public Communications]*, 56.

Verma, D. 2024. Impact of artificial intelligence on journalism: A comprehensive review of AI in journalism. *Journal of Communication and Management*, 3(02):150–156.

Wang, Y., M. Wang, H. Iqbal, G. Georgiev, J. Geng, and P. Nakov. 2024. Openfactcheck: A unified framework for factuality evaluation of LLMs. https://arxiv.org/abs/2405.05583. ArXiv preprint arXiv:2405.05583.