SYNTAKTIS: A Large-Language-Model-Backed Editing Interface for Supporting Ethical Journalism Practices

Leah J. Teichholtz
leahteichholtz@college.harvard.edu
Harvard University
Cambridge, Massachusetts, USA

Katy I. Gero
katy@g.harvard.edu
Harvard University
Cambridge, Massachusetts, USA

Elena L. Glassman
glassman@seas.harvard.edu
Harvard University
Cambridge, Massachusetts, USA

ABSTRACT

We present SYNTAKTIS, a novel interactive computational journalism tool backed by a large language model that automates the identification of several key ethical journalism problems. The interface aims to augment human journalists, not replace them, by offering optional feedback and revisions like an experienced editor. To do so, SYNTAKTIS is trained on the ethical journalism principles of the Society for Professional Journalists, which represent a gold standard taught in journalism schools and used by newspapers nationwide. We evaluate the efficacy of the interface through a user study with 14 student journalists and the quality of its output with a technical evaluation performed by two professional editors.

1 INTRODUCTION

In July 2023, the *New York Times* reported that Google was testing a new artificial intelligence product called Genesis that could help journalists write news articles. The tool aimed to serve as a "personal assistant" that would automate some tasks, the paper reported. Genesis met immediate resistance from journalists and news executives. Some found it "unsettling," adding that they believed it underestimated the amount of work and nuance needed to produce accurate and high-quality journalism [8].

This is just one example of the skepticism among newsrooms and the public over how and whether artificial agents should be used within the domain of journalism. It remains to be seen whether LLMs can be deployed for complex tasks unique to news journalism. Perhaps these machines are just not good enough for this task and it would be foolish to pursue them. Alternatively, LLMs could be proficient, but journalists may be reluctant to trust them with this work out of fear for their articles or job security.

There appear to be two opposing paradigms for using computational tools in news writing. On the one hand, LLMs can be used to generate entire articles and act as agentic assistants, as in the Genesis example. This threatens journalists' sense of agency and leads to lower-quality output, so most reporters find this use case uncompelling. On the other hand, mainstream AI tools used for the purely mechanical aspects of writing—like Grammarly [7]—are not useful for tasks more complex than knowing where to put a comma.

We introduce a browser-based tool called SYNTAKTIS, which aims to find a gray area between these existing uses of LLMs. Specifically, SYNTAKTIS attempts the complex task of editing for ethical journalism adherence, but it strives to do so in a way that maintains the full agency of reporters. The system provides in-line edits of words deemed to be sensational or biased as well as feedback and explanations of ethical concerns on issues that afflict a sentence. The interface also provides revisions and suggestions for new text though it never forces these on users. We sought to explore two research questions:

- *RQ*1: Can an LLM-backed ethical journalism tool help reduce the ethical journalism oversights in news articles?
- *RQ2*: Can an LLM-backed ethical journalism tool fit into existing editing processes and boost satisfaction but avoid authenticity and agency pitfalls that burden AI tools?

Participants in our user study offered feedback on the AI-generated edits and their experience using the interface. These users agreed or partially agreed with around 88% of sentence-level edits and 74% of word edits provided by SYNTAKTIS. Aside from enabling the detection of ethical errors, interviews completed during the user study suggest that the system maintained users' sense of agency. In addition, two longtime editors evaluated the output of the system on four real articles. The editors on average agreed or partially agreed with 75% of sentence edits and 72% of word edits. We also ran SYNTAKTIS on fabricated articles, and it identified most errors. This work makes several contributions:

- SYNTAKTIS, a novel LLM-backed interactive web interface that suggests edits to help journalists adhere to ethical journalism principles in news writing.
- Results from a technical evaluation of the output of SYNTAKTIS by two experienced editors and with fabricated pieces.
- Results from a user study with 14 student journalists at award-winning newspapers across the U.S. These participants brought their own pieces to the study, suggesting that the work could generalize broadly.

2 BACKGROUND AND RELATED WORK

To our knowledge, no academic work exists on the *use of computation for ethical journalism* purposes specifically, though the *ethical use of computation in journalism* has been discussed at some length.

2.1 Human-AI Collaborative Writing

Researchers have sought to apply the use of language models to writing and editing tasks. Several of these have focused on assistants that aid the mechanical aspects of writing and drafting prose. For example, Shi et al. developed a writing assistant to aid text completion, error checking, and keyword-to-sentence generation [14], while others created an assistant for spotting errors in technical reports [9]. Grammarly[1], Copy AI[2], Wordtune[3], and Writesonic[4] are just a few of the ever-expanding professional AI writing offerings.

[1]https://www.grammarly.com/, last accessed 3.13.2024
[2]https://www.copy.ai/, last accessed 3.13.2024
[3]https://www.wordtune.com, last accessed 3.13.2024
[4]https://writesonic.com/, last accessed 3.13.2024

Other HCI researchers have created tools that focus less on specific mechanics and more on broader creative ideation and writing processes. Zhang et al. created an LLM-backed writing assistant that helps users ideate, plan, and visually organize argumentative writing [19]. With Wordcraft, Yuan et al. studied the use of LLMs for the co-creation of short stories [18]. ABScribe, created and evaluated by Reza et al., builds on the idea of writing as an iterative process to enable users to create and visualize multiple variations of writing [13]. Meanwhile, Singh et al. developed and showed the efficacy of a multimodal machine learning support system for creative writing [15]. Finally, Dang et al. built a tool that generates real-time text summary annotations to enable users to reflect on and restructure their writing [3].

2.2 Computational Journalism

Journalism has long been a subject of interest for HCI researchers [20]. Computational or automated journalism is a growing field that involves any way that technology is used to discover, present, or monetize articles [2, 6]. Here, we broadly refer to computational journalism as encompassing any tools that support journalists throughout the article ideation, sourcing, and writing process as opposed to augmenting advertising or monetization workflows.

HCI methods have yielded useful and interesting tools for journalism. Petridis et al. developed AngleKindling, an LLM-based tool that helps journalists brainstorm angles for a story based on a press release [10]. Overview, created by Brehmer et al., allows for the systemic analysis and search of text documents for investigative journalism purposes [1]. Fulda et al. created TimelineCurator, which helps journalists temporally guide audiences through the automated generation of timelines [5]. Wang et al. built a tool to help journalists find sources from user-generated content [16].

2.3 Ethical Journalism

Journalistic codes of ethics aim to help reporters and editors decide what to do in particular situations. One code often seen as a gold standard is that of the Society of Professional Journalists (SPJ). The code starts with broad strokes language: "Seek Truth and Report It." "Minimize Harm." "Act Independently." "Be Accountable and Transparent." Within each section lies a series of bullet points that dictate the actual behavior associated with these categories. The ethical code used in this study relies on language that pertains specifically to writing articles, which is the focus of SYNTAKTIS.[5]

Though ethical codes in the field are widespread, research is mixed on their impact. A 1989 study by Pritchard and Morgan found that ethics codes had no direct influence on decisions made by journalists and, in effect, served more as ornamentation for news organizations [12]. Meanwhile, Fidalgo argues that journalism's ethical codes are ill-suited for the digital age, which is characterized by near immediacy of information and blurred lines between the citizen and the journalist [4]. Still, surveys of journalists globally suggest that organizational values have a larger influence on ethical code adherence than individual factors; this may indicate support for the use of ethical codes [11].

[5]https://www.spj.org/ethicscode.asp, last accessed 6.1.2024

3 SYSTEM OVERVIEW

In Ancient Greek literature, the word SYNTAKTIS, which means editor, is connected to the ideas of togetherness, arrangement, and the system of the world.[6] SYNTAKTIS captures effectively this interface's primary aspiration: working **together** with journalists to help them **edit** in alignment with a **system** of ethics. A video of the use of SYNTAKTIS is available at https://youtu.be/VjN3R00SqTA.

3.1 Prompting

We adopt multi-shot prompting in SYNTAKTIS, that is, providing a few examples of expected output given input to improve overall quality. We also force all output to be formatted in JSON, which is a toggle in OpenAI's API for the `gpt-4-0125-preview` model.

Initial versions of the interface attempted to coalesce all functionality in one prompt, but the LLM's performance was poor in these large prompts. Instead, recent work in chaining LLM input and output has shown that breaking complex tasks into smaller sub-tasks leads to better performance [17]. In SYNTAKTIS, we used three different prompts that targeted the individual paragraphs, full text, and headline of an article and analyzed it in the context of the ethical concerns outlined in sections 3.3 and 3.4.

3.2 System Architecture

The SYNTAKTIS landing page contains text input fields for a user's headline and article. The article text is capped at 10,000 characters to avoid exceeding the maximum token limit of the model, but more advanced models currently in development could likely enable the analysis of lengthier features and investigations.

Upon submission, the form data is sent to Flask. It is lightly preprocessed. For instance, words in quotes are removed before analysis to avoid flagging as sensational pointed language from an interviewee. Using the form data, Flask prompts `gpt-4-0125-preview`, which was released in January 2024. Flask with integrated Jinja2 uses the JSON data to generate a markup of the article, including the system's edits. The user is then free to interact with the edits and insert them into their article. Though the system adds highlights and underlines to the article, the output is a fully editable and interactive text editor: users may reformat text and add to the article as they find useful.

3.3 Word Identification

Avoiding the use of sensational and biased language is an important focus for SYNTAKTIS, which analyzes an article paragraph-by-paragraph and flags specific words or short phrases for potential ethical violations. When a user hovers over a highlighted area, the system provides replacement suggestions in a tooltip and specifies whether the word was judged to be sensational or biased.

The user can optionally click on one of these suggestions, which is automatically replaced inline and highlighted in green. In addition, the tooltip updates to show the previous word. This process is shown in Figure 1.

[6]https://logeion.uchicago.edu/ÏČÏŊÎ¡ÏĎÎśÎ¿ÎźÏĆ, last accessed 3.10.2024

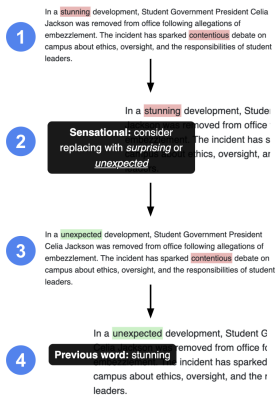

Figure 1: SYNTAKTIS word identification process.

3.4 Sentence Identification

We derive 18 different ethical journalism considerations from the SPJ's Code of Ethics, but some are infeasible for a computational tool (e.g. "Be cautious when making promises, but keep promises made") and others might violate norms about accuracy and agency (e.g. "Take responsibility for accuracy" is inadvisable given hallucinations of LLMs). We ultimately include seven ethical considerations in SYNTAKTIS: provide context, use anonymity sparingly and explain why it is granted, ask for comment, include key stakeholders, avoid stereotypes and generalizations, do not state beliefs as fact, and do not report in a way that distorts the truth.

When SYNTAKTIS identifies an instance of a sentence including a potential ethical oversight, it is thickly underlined in the error type's respective color. Clicking on the underline toggles the appearance of a suggestion box, which includes an additional sentence explanation for the ethical error and a revision of the sentence in a grey box. When a user hovers over the revision, a tooltip notifies them that they can click to insert the text. Clicking on the text replaces the original sentence with the revision in the article box; it also reformats the text, as in Figure 2.

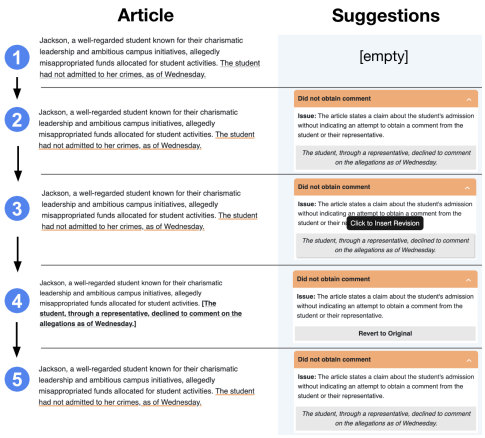

Figure 2: SYNTAKTIS sentence-level comments.

4 TECHNICAL EVALUATION

Experts. Two professional editors with experience in news reporting and editing were hired to evaluate the quality of SYNTAKTIS's output. The editors were each paid $50 to rate their agreement with edits provided by SYNTAKTIS on four real news articles.

The editors were provided with the four articles in a single Word document. Each identification from SYNTAKTIS was added as a comment on the document, and the editors were asked to respond to the comment with whether or not they agreed with the identification and why. We then tabulated the level of agreement with the identification. If an editor agreed that the sentence or word was justifiably identified, this was classified as agreement. Importantly, this was done based on the presence of an edit, not on the specific rewrite provided.

Editor 1 agreed with 75% of the sentence edits and disagreed with the rest, while Editor 2 agreed with 62.5% of the same edits, partially agreed with 12.5% of them, and disagreed with 25% of them. For word identifications, Editor 1 agreed with 48.3%, partially agreed with 10.3%, and disagreed with 41.4%. Editor 2 agreed with 79.3%, partially agreed with 6.9%, and disagreed with 13.8%. Notably, the editors did not always agree on whether the same edit was needed.

The editors also voiced concerns about the accuracy and usefulness of the rewrites in their comments. *E*2 agreed with a suggestion on the first article regarding including a key stakeholder, but wrote that they disagreed with the rewrite: "This is a secondhand quote. Be very careful with paraphrasing," they wrote. In the second article, *E*1 wrote that they "have a problem with the rewording because there is no evidence in the story that experts are suggesting this. It feels like the writer's opinion."

Fabricated Articles. Two brief pieces were written to include ethical concerns at the sentence and word level to see if SYNTAKTIS could spot the concerns. They were evaluated by SYNTAKTIS 10 times, which produced 10 different outputs on the same input for comparison. The first article, referred to as "Climate Protest," discusses a fictional protest by students calling on their school to divest from fossil fuels and voicing their opinions on the matter. The second article, dubbed "Student Government Ouster," reports on embezzlement allegations against the student body president.

SYNTAKTIS was overall effective on the two articles at identifying sentence-level ethical issues, detecting several of the issues in all 10 trials and overall 10 of 12 planted errors. It may show more variation when naming the specific concern.

A similar analysis was performed for biased and sensational word identifications in "Student Government Ouster." In all, 23 words were identified by the system, including 11 planted words. This means that 11 out of 12 planted words were flagged, but the system also flagged an additional 13 words that were not necessarily intended to cause concern. In other words, this feature has a high degree of false positives. Potentially, fewer than half of the flagged words would be real issues, meaning a user would need to dig through many highlights that would not be of use.

5 USER STUDY

This within-subjects study aimed to evaluate SYNTAKTIS from the perspective of a likely representative group of users: student journalists at independent university newspapers in the United States. 14 student journalists were recruited for the study, six of whom had never or almost never used LLM chatbots before.

5.1 Methodology

During the roughly 40-minute remote study, participants "edited" half of their article with SYNTAKTIS and half without any system support. The participants were screen- and audio-recorded. Participants were prompted throughout the study with questions about their thinking, and they also participated in a semi-structured interview after reviewing the entire article.

Participants were asked to bring an article that dealt with a sensitive topic or potential ethical issues. Users consented to participation, the use of their article as input for an AI model, and a video and audio recording. They additionally submitted an entry survey. In order to prime participants to think about ethical journalism, users read a one-page document that contained basic information about ethical journalism and avoiding sensational language.

The article was split in half by paragraphs. One half of the article was analyzed by SYNTAKTIS while the other half was not, and the half that saw SYNTAKTIS first was counterbalanced. During the control half, users did not have computationally generated edits and were asked to name ethical concerns that arose as they read.

When participants received SYNTAKTIS, they were provided a short demo. They were first given a few minutes to explore the interface freely. Then, they were asked to discuss their level of agreement or disagreement with each underlined sentence and the single-word suggestions; some participants moved into this stage without being explicitly told to do so. Finally, they were also asked to discuss if there were concerns that were not highlighted by the system. Users were also asked how they would remedy concerns.

After each half, users answered a survey that collected information on their satisfaction with the half they had reviewed and its adherence to ethical journalism standards. They also rated the system's functionality and usability and answered questions about their trust in SYNTAKTIS and its effect on their feelings of agency.

5.2 Results

10 of the 14 participants answered affirmatively that "all else equal, I would prefer editing with the system as opposed to not." None answered "no," and the remaining four participants said they were unsure. When asked to elaborate on the form, these four participants raised concerns about authenticity and "cheating," lack of accuracy and context, and data security.

5.2.1 Satisfaction with Piece for Editing. On average, participants were more satisfied with the half of the article reviewed with SYNTAKTIS than the control (4.36/5 vs. 3.86/5). Similarly, participants answered that, on average, they felt more prepared to formally edit the half of the article edited by SYNTAKTIS than the control. Participants were also asked to rate the level of "lingering ethical concerns" that they had for each half following their review. Responses following use of SYNTAKTIS responded on average 1.93/5, with none responding a 4 or a 5. Responses following the control were on average 2.36, with three users responding with a 4 or a 5.

Participants also commented on how using SYNTAKTIS felt like interacting with an editor. *P*4 remarked that the system fit "the model of an [editor]...this is their whole job." *P*7 said that the system was "not exactly like talking to an editor, but it pointed out the things that editors should." *P*14 noted that one edit was "exactly what my editors and I were discussing."

5.2.2 Ethical Concern Identification and Suggestions. On average, participants agreed with 63.6% of sentence suggestions, were in partial agreement 24.6% of the time, and disagreed with the suggestion 11.8% of the time. The number of comments made by the user during the control stage is also indicated, as shown, users generally did not make extensive comments on the control half.

Many participants said that the system aligned with concerns they had held or notes from their editors, and they said the system may have helped them think about more deeply about potential ethical pitfalls. For instance, *P*7 said it "helped me hone in on the specific bits of [missing] information." *P*10 said suggestions "align with my own criticisms of this article…"

On average, each participant agreed with the word identification 41.7% of the time, partially agreed or were unsure 32.7% of the time, and disagreed 25.6% of the time. Similar to the technical evaluation, there appears to be a high false positive rate with word identifications. Even with repeated prompting, it was challenging to get users to identify biased or sensational words and phrases during the control. 10 of the 14 user study participants did not identify any words of concern during the control half.

Users voiced mixed thoughts on the presence of the single-word identification. *P*5 noted flagged words "don't change the meaning of the article that much within the context of the story…They're good options to replace the words, but I think I just go ahead and keep them. That wouldn't be something that would flag my attention."

5.2.3 Changing Perceptions. Concerns about the dangers of AI tools arose during the study, which was unsurprising given that participants were generally skeptical about these tools. It's likely that using SYNTAKTIS may have shifted study participants' views on AI tools. Prior to the study, participants were overall neutral about the ability of AI to improve their journalism. But when asked whether tools like SYNTAKTIS could improve their journalism, participants strongly agreed. *P*3 said even if AI were to advance dramatically, they would "always prefer to have another person editing my stories, maybe even on top of AI, just because it's easier to talk through some edits…versus just putting full trust in an AI." *P*13 said it's important to use AI "as a tool, not to replace our job."

6 DISCUSSION

*Reducing Ethical Oversights (RQ1). RQ*1 asks whether an LLM-backed ethical journalism tool can ameliorate ethical oversights in news articles. Accuracy was measured in three ways: First, user study participants were asked to rate their agreement with the accuracy of edits on their own articles. Second, experts were asked to rate their agreement with edits left by SYNTAKTIS on real news articles. Finally, SYNTAKTIS's accuracy in evaluating fabricated articles with planted edits was measured. All evaluations offered evidence that SYNTAKTIS can powerfully diagnose many ethical errors.

SYNTAKTIS was imperfect—it made mistakes, mislabeled errors, and faced disagreement. It also had a relatively high false positive rate, specifically for word identifications. In this context, false positives are likely preferable to false negatives, as long as there are not so many false positives that the user loses faith in the system. Still, participants in the user study didn't give any indication that this had occurred. Generally, having a high rate of false positives may align with how real editors think about articles.

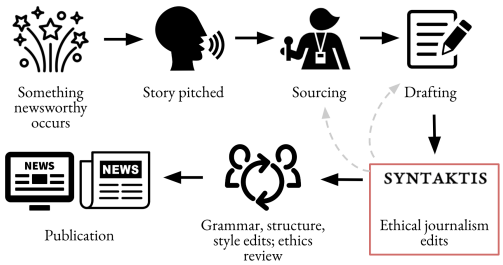

Figure 3: A workflow from event to article with SYNTAKTIS.

Journalism as a Process (RQ2). The second research question asks whether an LLM-backed system for ethical journalism can imitate the editing process for users while maintaining their agency. This implores us first to consider where a tool like SYNTAKTIS fits into the editing process, and how well it augments the editing process.

It is worth spending a moment here to discuss a typical news journalism workflow and where SYNTAKTIS might fit in. First, a newsworthy event occurs. Then, an editor pitches the story and prospective angles to a reporter. The reporter then interviews relevant stakeholders and does other data collection before drafting the story. In a typical workflow, the draft would then be reviewed by an editor. One could imagine inserting a step between drafting and editing—the use of SYNTAKTIS. If edits are substantial, a journalist can return to the sourcing phase and edit with a better-quality draft. In this revised workflow, time is saved by spotting errors earlier.

The user studies results suggested the feasibility of this vision, as participants said receiving edits with SYNTAKTIS helped them feel more prepared for formal editing. This workflow also demonstrates the interface's ability to maintain autonomy and control over the article—at no point did SYNTAKTIS replace a human.

7 LIMITATIONS

The methodology of the user study, which asked users to bring their own articles as opposed to providing them with one, came with significant advantages in terms of engaging users. At the same time, it also introduced other independent variables that may have distorted the data. Articles varied in length, topic, and ethical intensity. This meant that it was difficult to evaluate and compare results in an apples-to-apples manner.

Additionally, the user study split articles in half in a way that some participants noted felt unnatural—this is not a way that editing ever occurs. A better within-subjects experimental design might require participants to bring two article drafts; unfortunately, this is rarely a reality for busy journalists in the 24-hour news cycle. Future work could examine alternate experimental setups or move to a between-subjects design with a greater number of participants.

REFERENCES

[1] Matthew Brehmer, Stephen Ingram, Jonathan Stray, and Tamara Munzner. 2014. Overview: The Design, Adoption, and Analysis of a Visual Document Mining Tool for Investigative Journalists. 20, 12 (2014), 2271–2280. https://doi.org/10.1109/TVCG.2014.2346431

[2] Sarah Cohen, James Hamilton, and Fred Turner. 2011. Computational Journalism. 54 (2011), 66–71. https://doi.org/10.1145/2001269.2001288

[3] Hai Dang, Karim Benharrak, Florian Lehmann, and Daniel Buschek. 2022. Beyond Text Generation: Supporting Writers with Continuous Automatic Text Summaries. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology* (New York, NY, USA, 2022-10-28) *(UIST '22)*. Association for Computing Machinery, 1–13. https://doi.org/10.1145/3526113.3545672

[4] J. Fidalgo. 2013. Journalism is changing - and what about journalism ethics? https://www.semanticscholar.org/paper/Journalism-is-changing-and-what-about-journalism-Fidalgo/8e075cabd4687ef21cfb26f9b35f791372a45888

[5] Johanna Fulda, Matthew Brehmer, and Tamara Munzner. 2016. TimeLineCurator: Interactive Authoring of Visual Timelines from Unstructured Text. 22, 1 (2016), 300–309. https://doi.org/10.1109/TVCG.2015.2467531 Conference Name: IEEE Transactions on Visualization and Computer Graphics.

[6] Andreas Graefe. 2016. Guide to Automated Journalism. (2016). https://doi.org/10.7916/D80G3XDJ

[7] Hui-Wen Huang, Zehui Li, and Linda Taylor. 2020. The Effectiveness of Using Grammarly to Improve Students' Writing Skills. In *Proceedings of the 5th International Conference on Distance Education and Learning* (New York, NY, USA, 2020-08-28) *(ICDEL '20)*. Association for Computing Machinery, 122–127. https://doi.org/10.1145/3402569.3402594

[8] Benjamin Mullin and Nico Grant. 2023. Google Tests A.I. Tool That Is Able to Write News Articles. (2023). https://www.nytimes.com/2023/07/19/business/google-artificial-intelligence-news-articles.html

[9] Juanan Pereira and Mikel Alejo Barcina. 2019. A chatbot assistant for writing good quality technical reports. In *Proceedings of the Seventh International Conference on Technological Ecosystems for Enhancing Multiculturality* (New York, NY, USA, 2019-10-16) *(TEEM'19)*. Association for Computing Machinery, 59–64. https://doi.org/10.1145/3362789.3362798

[10] Savvas Petridis, Nicholas Diakopoulos, Kevin Crowston, Mark Hansen, Keren Henderson, Stan Jastrzebski, Jeffrey V Nickerson, and Lydia B Chilton. 2023. AngleKindling: Supporting Journalistic Angle Ideation with Large Language Models. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg Germany, 2023-04-19). ACM, 1–16. https://doi.org/10.1145/3544548.3580907

[11] Patrick Lee Plaisance, Elizabeth A. Skewes, and Thomas Hanitzsch. 2012. Ethical Orientations of Journalists Around the Globe: Implications From a Cross-National Survey. 39, 5 (2012), 641–661. https://doi.org/10.1177/0093650212450584 Publisher: SAGE Publications Inc.

[12] David Pritchard and Madelyn Peroni Morgan. 1989. Impact of Ethics Codes on Judgments by Journalists: A Natural Experiment. 66, 4 (1989), 934–941. https://doi.org/10.1177/107769908906600426 Publisher: SAGE Publications.

[13] Mohi Reza, Nathan Laundry, Ilya Musabirov, Peter Dushniku, Zhi Yuan "Michael" Yu, Kashish Mittal, Tovi Grossman, Michael Liut, Anastasia Kuzminykh, and Joseph Jay Williams. 2023. ABScribe: Rapid Exploration of Multiple Writing Variations in Human-AI Co-Writing Tasks using Large Language Models. https://doi.org/10.48550/arXiv.2310.00117 arXiv:2310.00117 [cs]

[14] Shuming Shi, Enbo Zhao, Duyu Tang, Yan Wang, Piji Li, Wei Bi, Haiyun Jiang, Guoping Huang, Leyang Cui, Xinting Huang, Cong Zhou, Yong Dai, and Dongyang Ma. 2022. Effidit: Your AI Writing Assistant. https://doi.org/10.48550/arXiv.2208.01815 arXiv:2208.01815 [cs]

[15] Nikhil Singh, Guillermo Bernal, Daria Savchenko, and Elena L. Glassman. 2023. Where to Hide a Stolen Elephant: Leaps in Creative Writing with Multimodal Machine Intelligence. 30, 5 (2023), 68:1–68:57. https://doi.org/10.1145/3511599

[16] Yixue Wang and Nicholas Diakopoulos. 2021. Journalistic Source Discovery: Supporting The Identification of News Sources in User Generated Content. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2021-05-07) *(CHI '21)*. Association for Computing Machinery, 1–18. https://doi.org/10.1145/3411764.3445266

[17] Tongshuang Wu, Michael Terry, and Carrie Jun Cai. 2022. AI Chains: Transparent and Controllable Human-AI Interaction by Chaining Large Language Model Prompts. In *CHI Conference on Human Factors in Computing Systems* (New Orleans LA USA, 2022-04-29). ACM, 1–22. https://doi.org/10.1145/3491102.3517582

[18] Ann Yuan, Andy Coenen, Emily Reif, and Daphne Ippolito. 2022. Wordcraft: Story Writing With Large Language Models. In *27th International Conference on Intelligent User Interfaces* (New York, NY, USA, 2022-03-22) *(IUI '22)*. Association for Computing Machinery, 841–852. https://doi.org/10.1145/3490099.3511105

[19] Zheng Zhang, Jie Gao, Ranjodh Singh Dhaliwal, and Toby Jia-Jun Li. 2023. VISAR: A Human-AI Argumentative Writing Assistant with Visual Programming and Rapid Draft Prototyping. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology* (New York, NY, USA, 2023-10-29) *(UIST '23)*. Association for Computing Machinery, 1–30. https://doi.org/10.1145/3586183.3606800

[20] Yujin Zhou and Zixian Zhou. 2021. Towards a Responsible Intelligent HCI for Journalism: A Systematic Review of Digital Journalism. In *Intelligent Human Computer Interaction* (Cham, 2021) *(Lecture Notes in Computer Science)*, Madhusudan Singh, Dae-Ki Kang, Jong-Ha Lee, Uma Shanker Tiwary, Dhananjay Singh, and Wan-Young Chung (Eds.). Springer International Publishing, 488–498. https://doi.org/10.1007/978-3-030-68449-5_47