AngleKindling: Supporting Journalistic Angle Ideation with Large Language Models

Savvas Petridis
sdp2137@columbia.edu
Columbia University
New York, New York, USA

Nicholas Diakopoulos
Northwestern University
Evanston, Illinois, USA
nad@northwestern.edu

Kevin Crowston
Syracuse University
Syracuse, New York, USA
crowston@g.syr.edu

Mark Hansen
Columbia University
New York, New York, USA
mh3287@columbia.edu

Keren Henderson
Syracuse University
Syracuse, New York, USA
khenders@syr.edu

Stan Jastrzebski
Syracuse University
Syracuse, New York, USA
sjastrze@syr.edu

Jeffrey V. Nickerson
Stevens Institute of Technology
Hoboken, New Jersey, USA
jnickers@stevens.edu

Lydia B. Chilton
Columbia University
New York, New York, USA
chilton@cs.columbia.edu

ABSTRACT

News media often leverage documents to find ideas for stories, while being critical of the frames and narratives present. Developing angles from a document such as a press release is a cognitively taxing process, in which journalists critically examine the implicit meaning of its claims. Informed by interviews with journalists, we developed AngleKindling, an interactive tool which employs the common sense reasoning of large language models to help journalists explore angles for reporting on a press release. In a study with 12 professional journalists, we show that participants found AngleKindling significantly more helpful and less mentally demanding to use for brainstorming ideas, compared to a prior journalistic angle ideation tool. AngleKindling helped journalists deeply engage with the press release and recognize angles that were useful for multiple types of stories. From our findings, we discuss how to help journalists customize and identify promising angles, and extending AngleKindling to other knowledge-work domains.

CCS CONCEPTS

• **Human-centered computing → Interactive systems and tools**; *HCI design and evaluation methods.*

KEYWORDS

Journalism, Brainstorming, Ideation, Large Language Models, Generative AI

CHI '23, April 23–28, 2023, Hamburg, Germany

ACM ISBN 978-1-4503-9421-5/23/04…$15.00
https://doi.org/10.1145/3544548.3580907

ACM Reference Format:
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 (CHI '23), April 23–28, 2023, Hamburg, Germany*. ACM, New York, NY, USA, 16 pages. https://doi.org/10.1145/3544548.3580907

1 INTRODUCTION

Journalists often write stories by carefully analyzing claims made in an interesting document for newsworthiness. These documents can be freely-shared, like press releases, private information leaks, like the Enron email dataset and Panama Papers, as well as public records accessed by Freedom of Information Act (FOIA) requests. Writing stories from a document is currently mentally taxing and requires a careful consideration of each claim's potential controversy and newsworthiness in order to brainstorm potential angles for stories. An angle is a *framing* of an event or document that "call[s] attention to some aspects of reality while obscuring other elements, which might lead audiences to have different reactions" [21]. Each angle forms a perspective from a few key claims of a document and sets the groundwork for developing a story. Then an angle is substantiated through interviews and information gathering from relevant resources. And by considering multiple angles for a document, journalists can make better decisions on what kind of story to write. But with shortages in newsrooms [58] and the abundance of these documents, journalists do not have the time to comprehensively explore multiple framings for each document.

Current computational tools for journalists predominantly support computational news discovery (CND) - the data-driven identification of potentially newsworthy information [16] [54]. CND tools are often built atop data streams like social media feeds and government websites to direct journalists to anomalous information that could lead to a story, such as a recent high-volume of posts or a recent document detailing a new algorithm the government is using. While these tools help direct journalists' attention to interesting information, they do not help them explore many story angles for that piece of information. From a three-month long co-design with four professional journalists, we learned that in the early stages of a story, an essential part of a journalist's process is to brainstorm multiple angles that can be substantiated into coherent and verifiable stories. For example, a new government policy or initiative might lead to many different controversies and negative outcomes, and currently, journalists rely on their expertise, which may be limited or biased, to consider these effects. In this work, we study how to help journalists explore multiple different angles, given an interesting document.

Large language models (LLMs) have shown great potential in many ideation tasks. Pre-trained on billions of text sources from the Internet, LLMs are a fundamental shift in natural language processing (NLP) [7]. They contain vast world-knowledge and are often able to generate fluent natural language text. With few to no examples, LLMs can reliably execute a number of complicated tasks, including summarization, information extraction, and ideation with remarkably fluent and accurate completions [10]. Within human-computer interaction, LLMs have been used for a variety of tasks, including helping science writers brainstorm ways to communicate their findings [24] and enabling creative writers to explore many ways of writing a story by generating character arcs [14]. In this work, we study if and how LLMs can help professional journalists brainstorm story ideas from a document.

While journalists write stories from many different documents, we focus on press releases. The journalists in the co-design emphasized that press releases, especially those released by government administrations, are a timely and important source of information for writing stories, which is in line with findings from prior work [42]. During a few of the co-design sessions, we observed the journalists as they brainstormed angles for press releases, and we formed four design goals for AngleKindling, our interactive web tool that supports angle exploration for a given press release. To help journalists cut through the fluff of the press release, AngleKindling summarizes it into a set of main points. To support angle ideation, AngleKindling employs an LLM to suggest potential controversies and negative outcomes. We provide these pessimistic angles since press releases are positively biased; during the co-design, the journalists mentioned that they are more interested in uncovering what the writers of the press release do not want revealed, namely potential negative impacts. Then, to help journalists verify these angles, AngleKindling links them to the source text, pointing to parts of the press release relevant to each angle. And finally, to provide context for each angle, AngleKindling provides a related news article as historical background.

To evaluate AngleKindling, we conducted a within-subjects user study with 12 professional journalists, comparing AngleKindling to the ideation features of INJECT [39], a recent creativity support tool for journalists that also supports angle brainstorming. INJECT utilizes more traditional natural language processing techniques to supply related articles and extracted entities relevant to the source text; these articles are grouped by their broader, overarching angles. Our findings show that participants found AngleKindling significantly more helpful for brainstorming ideas, while also requiring significantly less mental demand than INJECT.

To summarize, this work contributes the following:

- Four design goals for helping journalists explore angles for press releases, based on findings from our three-month long co-design with four professional journalists.
- AngleKindling, our interactive tool for exploring angles given a press release, which uses an LLM to generate numerous angles like controversies, facilitates trust by linking these angles to the source text, and provides historical background for each angle via a related news article.
- Findings from our evaluation, demonstrating that AngleKindling was perceived to be significantly more helpful for brainstorming story ideas, while requiring significantly less mental demand than the baseline. This was primarily due to AngleKindling (1) helping journalists recognize angles they had not considered, (2) providing angles that were useful for multiple types of stories, (3) helping journalists quickly and deeply engage with the press release, and (4) providing contextualized historical context.
- A discussion that highlights rich areas of future work, including enabling angle customization and prioritizing angles that are more promising than others. We also discuss how the techniques used in this paper can be applied to other domains like case law and academia, where LLMs can be used to explore how the decision made in one case might affect the outcomes of similar disputes and the ethical implications of academic papers, respectively.

2 RELATED WORK

2.1 Computational Journalism

The role of computation in journalism has predominantly been studied in two overarching categories: (1) understanding how technology shapes how consumers and audiences interact with and consume news media and (2) designing tools to improve journalists' capabilities and understanding how these tools affect their workflows [15]. Within the first category, past work has examined how news is shared [52], specifically the role search engines play in biasing the content we consume [55], as well as how users interact with and consume news [6] [22] [41]. Within the second category, numerous tools have been built that support a diverse set of journalistic tasks, including categorizing and understanding large document collections [9], identifying claims to be fact-checked [25] [1], and examining events and content on social media [61] [40] [19] [18] [37] [57]. This paper contributes to this growing body of research on supporting journalists' abilities with technology.

CND tools help orient journalists' attention to newsworthy information with algorithms [16] [54]. In more open-ended tasks, like exploring a large set of documents, CND tools often incorporate visualizations so that journalists interactively identify newsworthy information. For instance, *Overview* visualizes a hierarchy of document clusters in a tree-structure, to help journalists get a birds-eye view of the information within a corpus and identify branches that interested them [9]. Similarly, *Jigsaw* helped users explore document sets at the entity level, by visualizing their relationships across documents in a network [53]. As well as guiding users to newsworthy information via visualizations, other CND tools monitor data-streams for anomalies to find and draw attention to newsworthy data. For example, *CityBeat* monitors geotagged social media data to direct journalists to budding local events via abnormal spikes in posts [61]. Another system in this vein, *SRSR* (Seriously Rapid Source Review), directs journalists' attention to user-accounts on Twitter that might be useful sources for breaking news events [17]. Instead of social media data, *Algorithm Tips* monitors government websites for documents describing new algorithmic decision making systems [20]. To present promising leads to journalists, the system employs a crowd to rate each document on a few newsworthy categories. While these CND systems effectively direct journalists' attention to interesting documents, their support often ends there. Multiple angles can be spun from an interesting document, and in contrast to these CND systems, AngleKindling supports this process of brainstorming angles after an interesting document has been found.

The most comparable system to AngleKindling is INJECT, which also supports the creative process of brainstorming journalistic angles [39]. To help users discover angles for a particular topic, INJECT employs traditional natural language processing (NLP) techniques to provide suggestions of relevant people and articles with different types of angles: causal, quantifiable, and ramifications. The system also uses template-based "creative sparks", which are general suggestions that encourage journalists to consider how a related article's angle might be applied or related to the journalist's story. AngleKindling has similar design goals and also provides related articles and suggestions, but structures them differently. AngleKindling employs an LLM to generate potential *controversies*, *negative outcomes*, and *areas to investigate* from a press release; these are more specific suggestions directly tied to the text, compared to INJECT's sparks. Then to provide further context for these angles, AngleKindling connects each one to a related news article. Therefore, the two systems provide different types of creativity support: AngleKindling is *generative* and *specific*, synthesizing angles tailored to the press release, while INJECT is *associative* and *general*, providing references to previous, related angles and sparks that are more general. In this work, we conduct a user study comparing AngleKindling to INJECT's ideation features to understand (1) which kind of creativity support journalists prefer and (2) how LLM-based suggestions compare to providing articles with different angles in terms of helpfulness and cognitive load (3) and how to better design angle brainstorming systems.

2.2 Creativity Support with Large Language Models

Generative models are being successfully applied to support a number of creative tasks, including music composition [38], designing visual art [36] [35], and writing [24] [14] [47]. Large language models, in particular, are transforming a number of creative tasks. Trained on billions of documents, LLMs contain a vast amount of general world knowledge and can perform numerous NLP tasks without pre-training [10] [33]. LLMs have been used as open-ended collaborative writing tools. With *Wordcraft*, users can view multiple completions from an LLM as well as explore portions of text written in different styles [64]. LLMs have also been shown to be effective in more constrained contexts, including generating suggestions for science communication. Science writers found *Sparks* LLM-suggestions both interesting and useful as a means to understand a reader's perspective [24]. *BunCho*, also uses an LLM to generate titles and synopses from keywords [43]. Finally, LLMs are also currently being used to enable end-users to develop their own AI-infused applications in the form of LLM-chains, where the output of one LLM-step is fed into another [60] [59] [27]. In all of these applications, LLMs have been shown to be effective tools for increasing creativity and useful even when they provide unintended or incorrect outputs. In this work, we apply an LLM to generate angles for journalists, a context in which trust and verification is essential. To help journalists verify angles, we connect them back to the source text. In this work, we investigate if an LLM's common sense reasoning can help journalists think of angles they would not have otherwise.

2.3 Brainstorming and Ideation Tools

AngleKindling is very related to past work in brainstorming and ideation. To help inspire ideas, brainstorming tools often provide related information during the brainstorming session. For example, *InspirationWall* made real-time brainstorming discussions more productive by providing related concepts from a knowledge graph [2]. As well as related concepts, images are another great source of inspiration. Tools like *Idea Wall* [49] and *Idea Expander* [56] show that related images help increase the diversity of ideas generated in a brainstorm. In general, brainstorming tools have employed a number of different data-sources to provide related information, including knowledge graphs [2] [23] [5], word-associations [44] [29], images [49] [56], and crowds [3] [11] [63]. A key consideration for these systems is how "far" these related ideas should be. While providing slightly distant inspiration from the current set of ideas can positively impact the brainstorming session [5] [30], going too far, can often disrupt it [12]. In this work, we further examine how close inspiration should be to its source material, in the context of supporting angle ideation for journalists. We compare AngleKindling, which employs an LLM to generate angles directly stemming from a press release, to another creativity support tool for journalists which provides more associative and general creativity support, in the form of related articles.

In addition to providing related information, brainstorming tools will often organize these suggestions to help users (1) better understand the space of ideas and (2) think of more ideas. Often tools will cluster ideas based on their semantic similarity to present an overview of the idea space [51] [50] [13]. *IdeaHound* leverages individual crowd-workers' organization of ideas to generate an idea-map that helped users understand the diversity of ideas [51]. *RecipeScape* also constructs an overview of methods for cooking a dish, by automatically clustering recipes based on the order of their steps [13]. As well as clusters, other tools like *BlueSky* develop an ontology of ideas, to help highlight areas in the space that could use further brainstorming [26]. And finally, *IdeateRelate* provides multiple lists of related examples separated into higher-level categories, helping users more easily find the ideas closest to their own [62]. Beyond organizing ideas in relation to each other, a critical aspect of AngleKindling's organization is that it connects each angle to a relevant New York Times article, to provide historical background for the angle and give further context.

3 CO-DESIGN WITH PROFESSIONAL JOURNALISTS

We conducted a three-month co-design with four professional journalists (3 male and 1 female, average age = 38.25), with experience ranging from 2 to 28 years. The co-design's purpose was to develop a tool that would help journalists be more productive in writing stories. We met with the journalists for an hour each week, to discuss problems we could address and datasets to experiment with. From the earlier sessions of the co-design, we learned that exploring multiple angles was a crucial part of writing a good story, which is in line with past work [45]. An angle is a framing or perspective of a document that can be substantiated with extra context and information to create a more specific story idea. For example, given a newly proposed bill, one angle could be that a particular group of people might be unfairly affected. One story idea from this angle could be to analyze how related past laws have affected this group of people and how this proposed bill will add to or complicate these effects. Choosing an angle guides the remaining story writing process, including who to interview and what information to gather, and therefore it is critical that journalists explore multiple angles to choose a path that seems most promising.

After experimenting with a few datasets, we chose press releases as our source to support angle-exploration. We first experimented with social media posts made by prominent politicians, and used GPT-3, OpenAI's LLM, to ideate connections between a post and a number of angles, including health, technology, and economics. We applied GPT-3 to a post by Paul Gosar, an Arizona Representative who praised Florida's "Don't say gay bill", which forbids elementary school teachers from discussing sexual orientation with their students. We prompted GPT-3 to "List the potential implications of the actions described in the post on the American economy" and GPT-3 aptly connected Gosar's post to past events where companies had moved offices and jobs from states where anti-LGBTQ laws had been passed and concluded the same could happen in Florida. While this initial experimentation showed promise that GPT-3 could help journalists read between the lines of text to reveal its implications, we determined that social media posts contained too little text to extrapolate on for angles. Instead, the journalists pointed to press releases as a more applicable data source for angle exploration. Press releases contain more text and are often rife with positively biased claims for which LLMs can be used to identify implications. At the same time, writing stories from press releases is a very common task [8] [48] that is not well-supported by technology. Newsrooms are inundated with press releases, where individuals or teams are focused on churning out stories from them. Because of their prevalence in journalism and their claim-filled content, we focus on bettering angle exploration for press releases.

3.1 Formative Study: Brainstorming Angles from Press Releases

To understand how to best support angle exploration for press releases, we observed the journalists brainstorm angles for two press releases and interviewed them on their process. They were asked to imagine that their editor had handed them the press release to come up with a few potential story ideas. For each press release, they were given 15 minutes to come up with multiple, different story ideas, reflecting the time constraints of a newsroom. As they brainstormed, they were asked to record these angles in a separate document. We chose press releases distributed by New York City's mayor, Eric Adams, since the journalists lived in or nearby the city and would have the requisite background knowledge to identify the locations and individuals mentioned in the document. One press release (PR1) was about a new safety plan for the city's subway system and the other (PR2) was about plans for a new offshore wind hub to supply electricity for the city.

3.1.1 Findings: Design Goals. Each journalist started by carefully reading the press release. They noted how press releases are typically biased and filled with fluff, or less informative phrases that praise the government's actions. The journalists skim through this fluff to (1) quickly understand what the press release is addressing and (2) collect important information. This important information often included specific details pertaining to the plans described in the release. For the wind hub release, the journalists collected information on which companies were contracted to help with construction, the length of these contracts, the number of projected jobs, and other concrete information. These pieces of information were the foundation for the angles they brainstormed, but required a tiresome and potentially error-prone process of scanning the document for them. Therefore, our first design goal was to **summarize the press release into a set of main points, to help journalists quickly cut through the fluff and identify important details**.

From the claims they collected from the press release, the journalists ideated story angles, like potential controversies and negative outcomes. For the subway safety plan press release, P1 collected information on the number of police deployed at each station and their new role to remove homeless people from the subway at the end of each line. From this information, he wrote down two controversies: (1) "the increase in police presence may not reduce violence in the subway" and (2) "there might only be an increase in police brutality toward the homeless population". In addition to these potential controversies, he also wrote down questions for follow-up investigation, including "Have there been past subway plans and were they effective? Does increasing police presence normally reduce violence?" Similarly, for the wind hub press release, both P2 and P4 recognized that the city had employed a petroleum company to build the wind hub. P2 questioned how the petroleum company landed the contract as well as "how long they had been lobbying the city for this contract". P4 questioned if a petroleum company should be leading the city's green energy movement. Thinking of these controversies and questions was mentally demanding and therefore, **our second design goal was to provide angles that focus on elements of conflict and controversy**. Finally, the journalists brainstormed these angles directly from claims made in the press release, and thus **to facilitate trust in the angles we provide, our third design goal was to ground them in the source material**.

All four journalists emphasized the importance of getting historical background to either (1) think of new angles or (2) to get supporting information for the angles they brainstormed. While working on the subway safety plan release, P3 had questions including, "What have other cities done for subway safety plans?" and "How has New York's subway policy changed over the years?"

P3 stated that these are questions he would then answer by consulting past news articles written about New York's and other city's subway policies. He explained that by acquiring this historical background, new angles might appear, like "New York is trying the same, ineffective methods to mitigate subway violence" or "New York's new subway plan is radically different from that of other cities." As well as inspiring new angles, historical background can also provide supporting information for angles already brainstormed. For the wind hub press release, P2 became interested in the petroleum company's role in constructing the wind farm; he hypothesized that there could be tension from the local community and split opinions on the new hub. While he did not have time during the 15-minute time limit, he explained his next step would be to read past articles on this deal to see if they covered the local community's opinion before and find information to validate or refute his hypothesis. Therefore, **our last design goal was to provide relevant historical background to contextualize and spark new angles**.

Design goals. In summary, we formed four design goals for AngleKindling from the co-design:

D1: Cut through the fluff by summarizing the article into a set of main points.

D2: Provide angles focused on conflict and controversy to help journalists call in to question the positive bias of the press release and inspire story ideas.

D3: Facilitate trust by connecting the provided angles directly to the source text (the press release).

D4: Provide relevant historical background to provide further context, show what's been written, and inspire new angles.

4 ANGLEKINDLING

To address these design goals, we created AngleKindling: an interactive web tool that supports journalists in brainstorming angles, given a press release (Figure 1). AngleKindling displays the input press release on the right and the angle suggestions in the green sidebar on the left. The press release in this example is another by Eric Adams, announcing new zoning changes to improve New York's affordable housing and energy efficiency. To address D1, *cut through the fluff*, AngleKindling provides a list of the press release's main points (a_1), to help journalists skim the content quickly. To address D2, *provide angles*, AngleKindling provides a list of potential *controversies* (a_2) and *negative outcomes* (a_4) to offer an alternative perspective to the claims made in the press release, as well as *areas of investigation* (a_3) to offer questions the journalist might consider for inspiration. To address D3, *facilitate trust*, AngleKindling connects each angle and main point to five relevant portions of the press release. In this case, the user selected the second *controversy* (b_1): "The housing plan might not do enough to help those who are struggling to afford their rent or homeownership". AngleKindling then highlighted a relevant portion of the press release (b_2), which in this case, is a quote that directly opposes the *controversy*, claiming that the new zoning laws will improve the housing opportunities in less fortunate neighborhoods. Users can continue to skim through connected content with the related content button (b_3); the portion in-focus is highlighted yellow, while the rest are green. Finally, to fulfill D4, *provide relevant historical background*, a relevant article from The New York Times is retrieved for each angle (b_4). In this case, the article is from 2013, discussing how a past zoning measure had not improved conditions for low income New Yorkers, providing evidence for the *controversy* in (b_1) and contradicting the official's quote in (b_2). From here, the journalist can click on the blue arrow in (b_4) to read the article in full, and see if Eric Adam's new zoning plan proposes significant changes to past plans, or continue exploring other angles. Together, these features help journalists take an interesting source of information, like a press release, and explore multiple different story angles.

5 IMPLEMENTATION

AngleKindling is implemented in the Flask web-framework. To summarize the press release and generate the angles, AngleKindling employs GPT-3, OpenAI's large language model, via their API [1]. A central feature of AngleKindling is also connecting each angle to relevant sentences in the press release and a New York Times article. To connect content we embed the angles and press release content using Sentence-BERT [46], with the *all-mpnet-base-v2* pre-trained model in particular, via their API [2]. Finally, we use the New York Times API to link a relevant article to each angle [3]. In the following section, we describe how we use these tools to implement AngleKindling's core features.

5.1 Providing Angles and Main Points

Given the promise they showed in the co-design, we continued to use an LLM, specifically GPT-3, to fulfill D1 and D2 and generate angles for press releases. The press releases we collected were too long to fit in the input length of GPT-3. To generate angles across the entire document, we split the press release into a set of sections (Figure 2A) and generated angles for each section (Figure 2B). To split the document into sections, we separate the document into paragraphs, based on newlines and indents, and fit as many full paragraphs as we can, along with the prompt, in the input length of GPT-3. Initially, we used zero-shot prompts to ideate controversies for each of these sections. For each section, GPT-3 was prompted to "Create a list of controversies that could potentially arise from the following article section", without any training examples. The completions for the zero-shot prompt would sometimes produce a compelling result, but would mostly output generic, unhelpful controversies like "The plan will fail." At the same time, the completions were often phrased as facts, and instead, we wanted to hedge each controversy to present them in a less biased tone. Therefore, we switched to a few-shot prompt, for which the examples consisted of a press release section, paired with a list of controversies that the journalists thought of from the formative study (Figure 2B). The resulting angles, like "The plan could lead to more traffic and congestion in New York City" were often more specific and hedged to emphasize that they were *possible* controversies instead of facts. Finally, these resulting angles are selected from a single LLM-prompt run.

Extracting the main points of the press release is done similarly to the angles, but involved an extra challenge in removing the press

[1] https://openai.com/api/
[2] https://www.sbert.net/
[3] https://developer.nytimes.com/

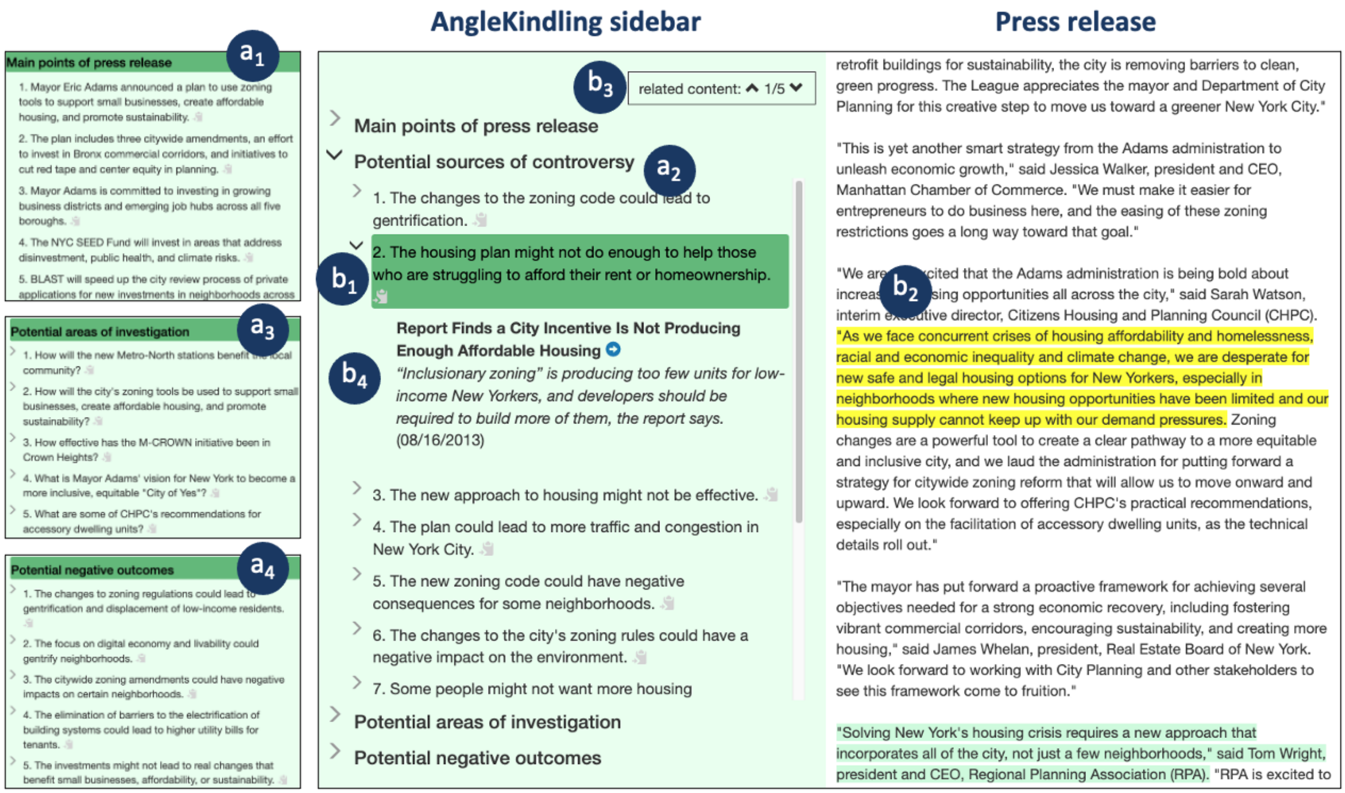

Figure 1: AngleKindling's interface displays the press release on the right and the article's main points (a_1) along with angle suggestions in the green sidebar on the left. The angle suggestions include potential *controversies* (a_2), *areas of investigation* which are questions to consider (a_2), and *negative outcomes* (a_4) that could arise. To help users trust these angles, they can select them (b_1) to view related content from the press release (b_2), and they can skim through up to five pieces of text with the related content button (b_3). Finally, each angle is connected to a New York Times article from the past decade (starting in 2012) to provide historical background (b_4). The title, lead paragraph, and publication date are provided for the article, as well as a link to the article itself, via the blue arrow.

release fluff within each point. From the same sections we used to generate the controversies, we also generate the main points, using another few-shot LLM prompt. However, each main point tended to include superfluous information that only served to further the document's positive bias. For instance, from the offshore wind press release in the formative study, one main point was that "New York City Mayor Eric Adams today announced an agreement that will transform the city-owned South Brooklyn Marine Terminal (SBMT) into one of the largest offshore wind port facilities in the nation" To simplify this point, we use another few-shot LLM prompt to rewrite it with fewer words, generating the simpler, less-biased sentence: "Mayor Adams announced that the South Brooklyn Marine Terminal will be turned into an offshore wind port." Like the controversies, we run the LLM-prompt once and use that generation as the main points. With this extra step, we are able to provide a summary that is easier to read and better cuts through the fluff.

5.2 Connecting Angles to the Source Text and Historical Background

While these main points and angles might be accurate and inspire ideas, they are difficult for journalists to trust without explicitly tying them back to the source material. To help facilitate this trust, we identify each angle's top five most related sentences in the press release. To do so, we compare the similarity between each angle to each sentence in the press release. For each angle we compute a vector, using Sentence-BERT. Next, we split the press release into sentences using spaCy's[4] built-in sentence segmentation; each sentence is then embedded, also with Sentence-BERT. Finally, we compute the cosine-similarity between each angle to each sentence, and the top five sentences are selected to be highlighted by the related content button (Figure 1b_3). By explicitly connecting each angle and main point to the source text, we help journalists quickly verify their relevance.

As well as connecting each angle to the source text, another crucial feature of AngleKindling is bringing historical background

[4]https://spacy.io/

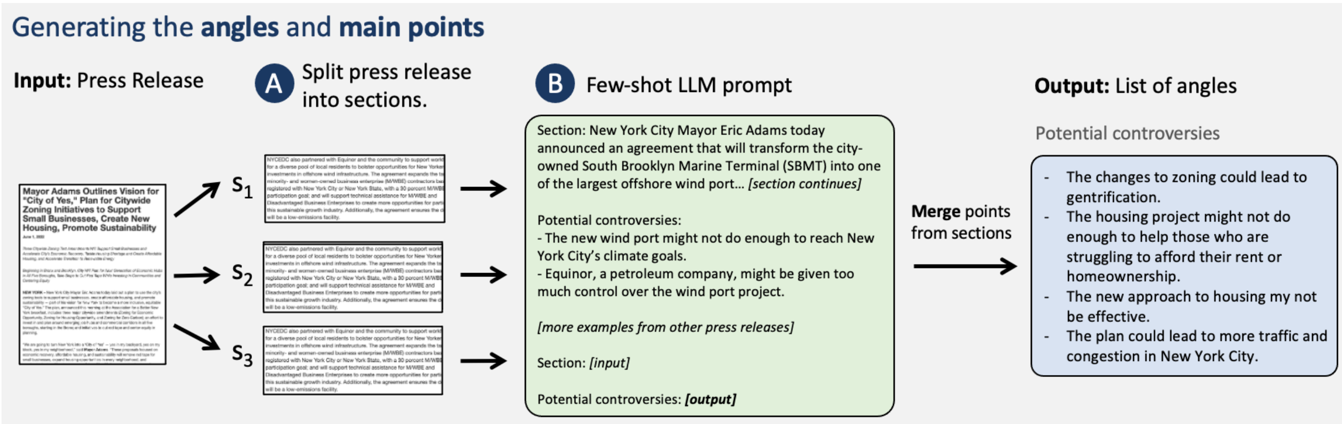

Figure 2: To generate the angles and main points, AngleKindling first splits the press release into a set of sections, to fit the input length of the LLM (A). Each section is then inputted to a set of four LLM prompts, to (1) extract the main points of the section (2) ideate potential *controversies,* (3) identify *areas to investigate,* and (4) ideate potential *negative outcomes* (B). Each LLM prompt is few-shot and contains three examples of converting a section into a set of main points or angles. The examples are taken from the angles thought of by the journalists in the formative study. Finally, the angles ideated from each section are then merged together into a single list.

by connecting each angle to a past news article. We use The New York Times (NYT) as our source of news articles, as it is (1) a reputable and exemplary news source trusted by journalists and (2) likely to cover the important problems and plans that pertain to New York City. To connect each angle to a news article, we first collect a set of relevant NYT articles for the press release. To do so, we extract the top five most relevant keywords from the press release, once again with a few-shot LLM-prompt. Each keyword is then used to query New York Times articles from the past decade, using their developer API. Through this process we normally collect approximately 300 relevant articles. For each relevant article, we concatenate its headline and first paragraph to compute an embedding using Sentence-BERT. Often the first paragraph of a news article will convey the most important facts of the story, which along with headline, can be used as the representative material for the article. We then compute the cosine-similarity of these headline embeddings with each angle embedding, and choose the highest scoring article to use as historical background. By doing so, we help journalists gather context for each angle through relevant historical knowledge.

6 EVALUATION

To understand how AngleKindling may help journalists brainstorm story ideas, we conducted a within-subjects study, comparing AngleKindling to the ideation features of INJECT, a comparable creativity support tool for journalists. To understand what participants liked and disliked about these systems we (1) include a questionnaire to get quantitative measures for each tool's features and helpfulness as well as (2) conduct a semi-structured interview to get qualitative insights on participants' preferences.

6.1 INJECT

INJECT is a creativity support tool created to help journalists write stories faster by helping them discover creative angles. INJECT's use case is a bit different from that of AngleKindling. Implemented as a Google Docs Add-on sidebar, INJECT provides creativity support for the story the journalist is currently writing. Journalists highlight text in their story, from which INJECT extracts keywords to retrieve associated news articles. To help journalists apply these articles, INJECT includes creative "sparks", which give general advice on how to use these articles, like "Take this story. What does the backstory inspire?". While INJECT provides *general* creativity support for what the journalist is currently writing, AngleKindling provides *specific* creativity support for a given source, by directly generating angles from its content. While the use-cases of the two systems are different, their aim is to support angle-ideation by helping journalists expand content, like what they have written already, or an interesting document source, to help them see this content from a new perspective.

In its evaluation, INJECT's ideation-features were shown to be quite helpful for journalists. Deployed in multiple news outlets, INJECT was incorporated in the workflows of professional journalists and used to develop multiple, published stories. As well as becoming a staple in journalists' workflows, INJECT also improved their writing efficiency, helping journalists think of new angles for their stories faster, "often in less than 3 minutes for each story" [39]. Thus, INJECT's general ideation-features were shown to be quite powerful, and so we wanted to compare AngleKindling's specific creativity support to INJECT's as a baseline.

We selected INJECT's most relevant features and compare these to AngleKindling, as opposed to INJECT itself. The original INJECT includes six sources of creativity support, of which we include four (all based on prior news articles): (1) **Quantifiable**: articles that contain quantified information, such as actual numbers, and

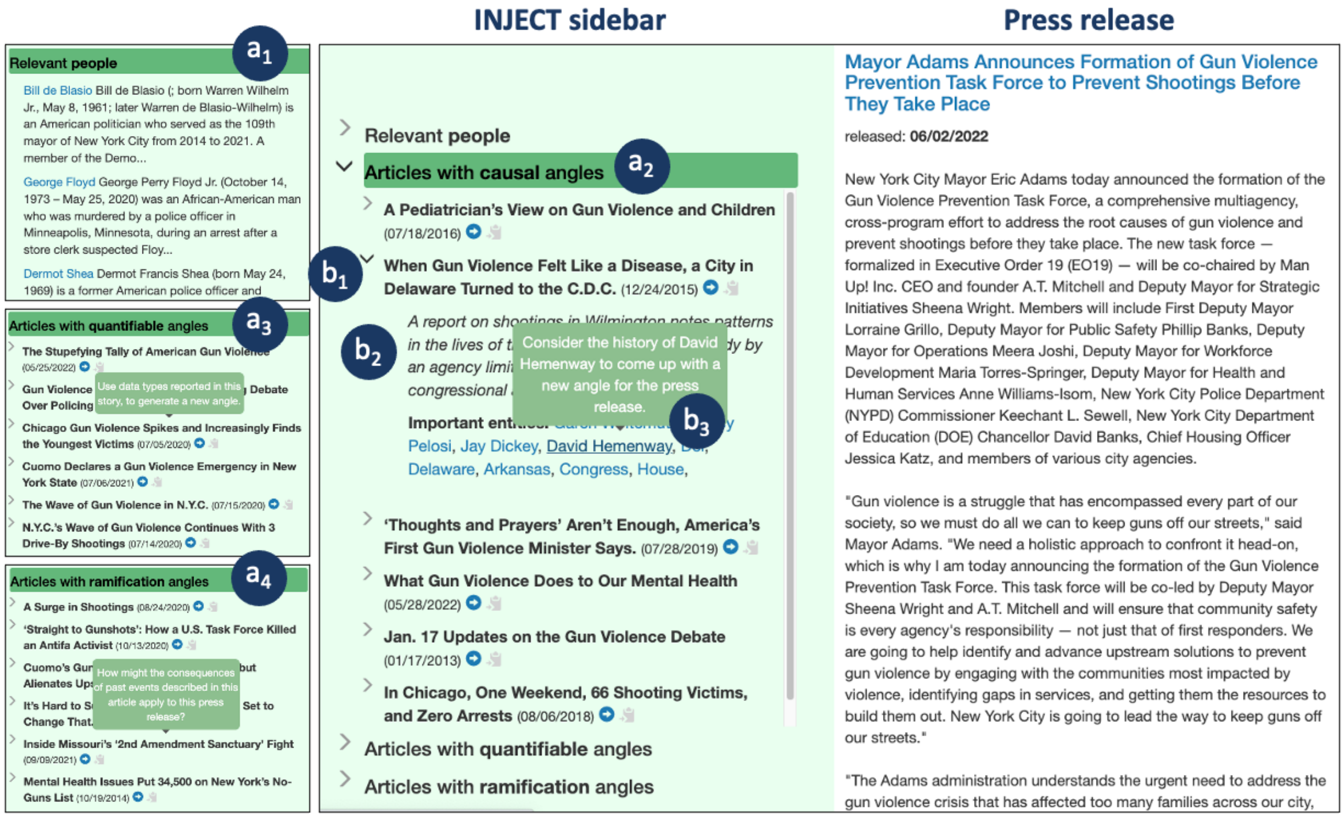

Figure 3: INJECT's interface incorporates four types of information sources: relevant people (a_1), articles with *causal* angles (a_2), articles with *quantifiable* angles (a_3), and articles with *ramification* angles (a_4). Each article (b_1) is clickable to reveal its first paragraph (b_2), as well as a list of its extracted entities (people, places, organizations, and events) that are linked to their corresponding Wikipedia pages. Each entity can also be hovered over to reveal a inspirational "spark" (b_3). These sparks also appear when a user hovers over an article title (a_4).

keywords like *Sterling* and *population*, (2) **People**: information on individuals (from Wikipedia) extracted from related news articles, (3) **Causal**: articles that discuss the background or causes of a story, identified through keywords like *cause*, *impact* and *studies*, and (4) **Ramifications**: articles that discuss the future consequences of a story, identified through keywords like *outcome*, *consequence*, and *aftermath*. INJECT's two remaining features include associated (5) news comics and (6) data visualizations, but these were less related features to those provided by AngleKindling.

To provide the news articles for our selected features of INJECT (Causal, Quantifiable, Ramifications, People), we use the same dataset of articles pulled from The New York Times that we collected for AngleKindling, described in Section 5.2. INJECT originally has search functionality, but in our case, we assume the articles have already been searched for, using keywords from the press release. To minimize the visual difference of the two interfaces, we incorporate INJECT's articles as drop-downs for each category: *people* (Figure 3a_1), *causal* (Figure 3a_2), *quantifiable* (Figure 3a_3), and *ramifications* (Figure 3a_4). The *people* are extracted from the articles shown in the other categories and sorted by frequency. They are also linked to their Wikipedia pages. Next, the articles are assigned to a category (*causal*, *quantifiable*, and *ramifications*) based on a set of pertinent keywords for each one. When users select a category, they can view its articles (Figure 3b_1), along with each article's publication date, a link to its page, its first paragraph, as well as its extracted entities: people, places, organizations, and events. Each entity is linked to its corresponding Wikipedia page and like INJECT, includes a hover-over "spark" related to its category (Figure 3b_3). These sparks also exist for the article headlines (Figure 3a_3), and are generated using templates provided from the original paper. With this selection of INJECT's features, we can compare its associative and general creativity support to AngleKindling's more specific creativity support.

6.2 Procedure

The general outline of the study was the following: (1) participants were first interviewed on their journalism background and experience, (2) they then used AngleKindling and INJECT to brainstorm story angles for two press releases by New York City's mayor, (3) after brainstorming with each tool, they filled out a questionnaire

Metrics (Both Conditions)	Statement (7-point Likert scale)
Helpfulness	The system as a whole was helpful for coming up with story ideas.
Pursuable Angles	I would pursue some of the angles from this system.
Mental Demand	Coming up with story ideas was mentally taxing with this system.

AngleKindling Metrics	Statement (7-point Likert scale)
Main Points	The main points were helpful for skimming the press release.
Related Content	The "related content" button helped me find relevant information in the press release.
Controversies	The controversies were helpful for coming up with story ideas.
Areas to Investigate	The areas to investigate were helpful for coming up with story ideas.
Negative Outcomes	The negative outcomes were helpful for coming up with story ideas.
Historical Background	The articles were helpful for coming up with story ideas.

INJECT Metrics	Statement (7-point Likert scale)
People	The relevant people provided were helpful for coming up with story ideas.
Causal	The articles with causal angles were helpful for coming up with story ideas.
Quantifiable	The articles with quantifiable angles were helpful for coming up with story ideas.
Ramifications	The articles with ramification angles were helpful for coming up with story ideas.
Sparks	The hover-over creative sparks were helpful for coming up with story ideas.

Table 1: Post-task questionnaire filled out by participants after using either AngleKindling or INJECT. For both systems, participants were asked to rate its *Helpfulness, Pursuable Angles* and requisite *Mental Demand.* Each system also had its own specific statements for rating each of its features, to gauge what was most helpful of each tool.

rating each tool's features and their experience coming up with ideas, (4) in a semi-structured interview, they were then asked a series of questions on their preferences and thoughts on each tool.

In the experiment phase of the study, participants were randomly assigned to a condition that determined which tool and press release they would brainstorm story ideas with first. Tool and press release order were counter-balanced to prevent a learning effect. Participants were asked to imagine that their editor had assigned them the press release and asked them to come up with many different story ideas for the press release. Before using each tool, they were shown a video demonstrating its features, using the offshore wind press release as an example. Also, since both tools used New York Times articles, participants were given login information for the publication if they did not have a subscription. After they felt they understood each tool's features, they were then given 15 minutes to brainstorm story ideas for the press release. From the co-design, we found that this time-limit was reasonable and reflective of the time constraints that many journalists face in practice at daily news publications. Participants recorded their story ideas in a document and were encouraged to explain their process and reasoning as they came up with ideas. We described "story ideas" to participants loosely as questions or lines of thought they were genuinely interested in pursuing. After coming up with story ideas with each tool, participants were asked to fill out a questionnaire (Table 1) to understood how each tool and its features helped them brainstorm ideas. And once they brainstormed ideas for both press releases, they were asked a series of questions that probed their preference of each system, how each tool did and did not help them, and how they can be improved.

	AngleKindling	INJECT	p-value
Helpfulness	6.17 (0.99)	3.92 (1.38)	**<.05**
Pursuable Angles	6.33 (0.75)	4.5 (2.25)	.058
Mental Demand	1.83 (0.9)	3.42 (1.89)	**<.05**

Table 2: Comparison of AngleKindling and INJECT across the three categories from the questionnaire. We conducted three paired-sample Wilcoxon tests with Bonferroni correction, and found that AngleKindling was perceived to be (1) significantly more *helpful* and (2) significantly less *mentally demanding* to use for brainstorming story ideas. Average scores are shown with standard deviation in parenthesis. Significant p-values are bolded.

6.3 Participants

We recruited 12 professional journalists (average age = 37, 3 male and 9 female, experience in the field ranging from 5 to 29 years) via e-mail and social media calls for participation. None of the journalists in our co-design took part in this study. Eligible participants included journalists that work in any medium, including digital publications, newspapers, magazines, radio or TV. Since the press releases were in English and from New York City, we required participants to be English speakers and based in the United States. Also, part of the selection criteria was that participants must have written stories from press releases in the past. The interviews were conducted remotely, and participants had to have a computer with Google Chrome. Participants were compensated $30 for up to 60 minutes of their time.

AngleKindling Feature Ratings	
Main Points	6.17 (1.14)
Controversies	5.92 (1.26)
Negative Outcomes	5.75 (1.42)
Related Content	5.67 (1.18)
Areas to Investigate	5.33 (1.75)
Historical Background	4.67 (1.31)

INJECT Feature Ratings	
Quantifiable	4.83 (1.62)
Causal	3.83 (1.07)
Ramifications	3.75 (1.83)
Sparks	3.08 (2.1)
People	2.5 (1.61)

Table 3: The questionnaire results for each condition's features. AngleKindling's highest rated features were the *Main Points* and potential *Controversies*. The *Main Points* along with the *Related Content* helped users deeply engage with and understand the press release quickly, while the *Controversies* provided many, different ideas for stories. INJECT's highest rated feature were the *Quantifiable* articles which many journalists appreciated as a source of data and ideas for incorporating analysis in their stories. Average scores are shown with standard deviation in parenthesis.

7 RESULTS

From the exit-interviews, all 12 participants preferred AngleKindling to INJECT for brainstorming story ideas. Since the study was within-subjects and the questionnaire involved ordinal data, we conducted three paired sample Wilcoxon tests with Bonferroni correction to compare the two systems' *helpfulness*, how *pursuable* their angles were, and their requisite *mental demand*. We found that **AngleKindling was perceived to be significantly more helpful for coming up with story ideas** ($W = 55$, $Z = 2.96$, $p < .05$), scoring on average 6.17 (std = 0.99) on the questionnaire, while INJECT scored 3.92 (1.38) (Table 2). Furthermore, while also more helpful for brainstorming ideas, **AngleKindling also required significantly less mental demand** ($W = 0$, $Z = -2.74$, $p < .05$), scoring on average 1.83 (0.9) compared to INJECT's 3.42 (1.89). Finally, while not significant, participants on average also rated AngleKindling's angles as more pursuable (avg = 6.33, std = 0.75) than those by INJECT (avg = 4.5, std = 2.25).

In the following subsections we provide greater context to these results and illustrate that AngleKindling was more helpful because it (1) reduced the cognitive load of brainstorming angles with *specific angles that easily inspired next steps*, (2) provided angles that were useful for *multiple*, *different* types of stories, (3) helped journalists *quickly* and *deeply* engage with the press release, and (4) incorporated *contextualized* historical background.

7.1 AngleKindling reduced the cognitive load of brainstorming angles by providing specific angles that easily inspired next steps.

The angles produced by GPT-3 in AngleKindling were specific and easy to imagine potential stories and follow-up research with, reducing the cognitive load of brainstorming. For example, while working on the gun violence release, P2 found the following *controversy* promising: "There could be resistance from community members who don't want more police presence in their neighborhoods". This potential *controversy*, spurred P2 to consider (1) what specific actions the task force will take and which communities they will impact (2) how those communities have been historically impacted by police and gun violence, and (3) interviewing these communities on this task force. For the zoning press release, P9 was surprised by the *controversy* that "The plan could lead to more traffic and congestion in New York City." She had not considered this effect and became interested in interviewing city-planning experts about the new train lines proposed by the plan. Overall, AngleKindling's angles were specific and helped journalists easily consider what information to gather, which questions to ask, and who to interview.

However, some of AngleKindling's angles were too generic to imagine specific next steps. For example, for the *controversy*: "The plan does not do enough to address the housing crisis," P11 explained that this angle could be "made by anyone about anything". The angle is too broad to inspire a specific line of reasoning or next steps. These generic angles did not overly inhibit participants however; they were quickly able to scan each set of angles for anything specific and interesting. Overall, the participants appreciated being able to recognize interesting angles instead of coming up with them on their own. P8 described AngleKindling as "proactive", helping her to immediately "see how I could write several stories from this one press release".

Meanwhile, with INJECT, participants had to work harder to get to specific angles and think of next steps. Their process involved assessing angles other journalists had used in the news articles and determining if they could be applied to the press release. As P5 describes, "This one [INJECT] is more: think about what other people did on a similar story and apply it here." For example, for the article "New Jersey Town says 'No Thanks' to Development", P5 explained she might read this story to find out the reasons why the residents of this town wanted less development and see if they're applicable to New York residents where the zoning changes were being made. Then after this work, she might have a better idea of what to ask New York residents in an interview. Thinking of angles this way was more mentally demanding and likely led to the significant difference in rating (Table 2). To get to the same point as AngleKindling's angles, participants had to skim through articles, collect information, and mentally reason if it was applicable to the press release. The hover-over sparks did little to make this process easier, as shown by their low average rating of 3.08 (Table 3). Participants found sparks like "Make your angle more similar to the causal angle in this story" too high-level to be helpful. Overall, coming up with story ideas with INJECT involved a few more mentally taxing steps, while AngleKindling preemptively processed the press release to provide actionable and specific angles that easily inspired next steps.

7.2 AngleKindling's different angles were useful for multiple types of stories.

Participants found that AngleKindling's angles to be useful for multiple different stories, including (1) day-of stories, (2) next-day or week-long investigations, and finally (3) months-long retrospective stories. Multiple participants, including P4, P7, P9, and P12 found the *areas to investigate* particularly useful for **day-of stories**: briefer pieces that aim to summarize the key takeaways of the press release. The *areas to investigate* often included questions that P4 described as "aiming to clarify" the press release and useful for gathering information, such as the following: "What types of services and programs will be offered through this task force?". However, these kinds of stories were less exciting to many of the journalists who instead, preferred investigations that probe what the administration "does not want revealed", as P4 stated. This reasoning likely led to *areas to investigate* have the lowest average usefulness of the GPT-3 completions incorporated in AngleKindling (Table 3). Meanwhile, more investigative story ideas stemmed from the potential sources of *controversy* and *negative outcomes*. For example, P12 pointed to the *controversy* that "AT Mitchell may not be qualified to lead the task force" as a potential **next-day or week long story**. She explained that, over the course of a few days, she would research the communities that would be most affected by the new gun violence prevention measures Mayor Adams enacted and then interview organizations or prominent members of those communities to get their take on AT Mitchell. Finally, P8 pointed to a *negative outcome* that could potentially become a **months-long retrospective story**: "The task force could be used to unfairly target communities of color". She explained that after a few months after the press release was distributed, she might gather some data on who was arrested and where police were being stationed to gauge if this *negative outcome* had come into fruition. Overall, AngleKindling's different angles lent themselves to multiple types of stories, from shorter, summarization pieces to longer, deeper investigations.

7.3 AngleKindling helped journalists deeply engage with the press release quickly.

Many participants noted that the press releases from Mayor Adams' administration were complex and filled with unnecessary details that diverted their attention from important information. The main points (D1) and the highlighted related content (D3), which participants rated on average 6.17 and 5.67 respectively (Table 3), helped them quickly skim and understand the press release, despite this distracting fluff. Participants predominantly used the main points not to replace the press release but to supplement their reading of it. One common strategy they employed was to use the main points as a reading guide: they first scanned the main points to get a (1) high level view of the claims and (2) a quick idea of the information they found interesting, then read the press release in its entirety. As well as guiding their reading, participants also used the main points as a quick reminder of the press release's content as they thought of story ideas. After reading the press release, and throughout his brainstorming process, P10 would reread the main points to regain a "holistic view" of the press release as he assessed potential sources of *controversy* and *negative outcomes*. By doing so, he could better contextualize and make sense of each potential angle. However, the main points were not perfect. P12, P5, and P8 mentioned that the main points missed information that they were very interested in from the press release, particularly the specific implementation details and statistics included in the document. These concrete details are very useful for critically examining the feasibility of the plans mentioned in press releases. Thus, the main points helped participants quickly understand the press release, but at the same time, can be improved to prioritize the statistics mentioned in the document.

Highlighting related content was critical to helping journalists trust both the main points and generated angles. As P2 explains, "The highlighted text is useful. It takes me there right to it... I would not trust these main points without the highlighted text." The related-content button and highlighted text helped the participants quickly verify the veracity of each main point, and without this feature, they would be concerned that the main points might be erroneous or misleading. The related content button also helped journalists acquire evidence to better understand a potential *controversy's* source. While working on the gun violence press release, P12 came upon a *controversy* that was completely unexpected: "There might be infighting among the various agencies involved". She did not immediately understand why this *controversy* might be related, so she used the related content button to scrub the press release and was taken to a claim in the text that explained the new gun violence task force would work with multiple agencies, including the departments of health, social services, and housing. Being able to verify these *controversies* enabled journalists to find interesting information they had not previously considered in the press release.

7.4 AngleKindling provided *contextualized* historical background, which helped with brainstorming story ideas.

Connecting a prior news article to an angle helped journalists better understand how the article was related and how it could be applied to inspire new story ideas. For example, for the zoning press release, P4 selected a potential *negative outcome* that stated, "The increased housing opportunities might not be affordable for low and middle-income New Yorkers." The connected news article was entitled "Some 'Affordable' Units Too Costly, Report Says" and detailed how new affordable homes being built in the Bronx required household incomes above the median in New York City. The combination of this angle and news article inspired the potential idea of comparing the new plan with this past attempt to create affordable housing, to answer questions like: Does Eric Adam's plan avoid the pitfalls of past plans? Are these more empty promises? However, sometimes articles did not provide useful background because they were tenuously connected to the angle. For example, in the gun violence task force press release, the *negative outcome*: "The recommendations of the task force might not be implemented properly" was connected to an article about who is on U.S. Coronavirus Task Force. P2 stated this could be potentially interesting toward a broad story on the general effectiveness of task forces, but ultimately found this article less useful because it was describing a federal task force instead of a city task force, created for a very different problem. Overall, when the news articles were closely related to the angle, participants were able to get relevant background information that sparked new ideas.

Meanwhile, in the INJECT condition, the separation of news articles into causal, quantifiable, and ramification angles was not very conducive to coming up with story ideas. Participants had trouble discerning why a certain article belonged to one of the categories, especially the causal and ramification groups. P4 stated, "I did not really get the causal or ramification angles. This information didn't come through the article headlines". He was unsure of how the article "A Pediatrician's View on Gun Violence and Children" belonged to the causal category; it was not immediately clear what background or causes this article would reference. However, most participants appreciated the quantifiable category, aligning with the findings from INJECT's own evaluation [39]. Among INJECT's features, the quantifiable articles were the highest rated, receiving an average score of 4.83, compared to 3.83 for the causal articles and 3.75 for ramifications articles (Table 3). The articles within the quantifiable category, were more explainable, often containing a statistic in their headline or lead paragraph, like: "It has been nearly a quarter century since New York City experienced as much gun violence in the month of June as it has seen this year." The quantifiable articles also provided inspiration on potential datasets to use or analyses to conduct for the press release. P11 stated, "I really like the quantifiable angles, they include numbers and even trends that help give me context for my story." Finally, the participants also appreciated that the articles appeared together in longer lists, which helped provide great coverage of the topic as a whole. P9 explained, "[INJECT] is a bit broader. It helps me better understand the topic...this a great tool for background information." INJECT's list-organization, while not immediately useful for brainstorming ideas, helped participants better learn about the topic as a whole.

Finally, for both systems, participants wanted more contextual information, beyond historical news articles. Many participants mentioned that their goal is to go from the source material to interviewing relevant people and organizations as fast as possible. P4 explained, "The best story ideas will come from people who are smarter than me on the topic." He wanted both tools to go beyond providing angles and provide *local* organizations, leaders, and experts to interview. From these interviews, journalists can identify the most important questions to answer in a story. While INJECT extracted people and organizations for its related articles, often the extracted individuals were too famous to easily interview or not related or local enough to the press release. As well as individuals to interview, P6 who has a background in law, wanted excerpts of relevant laws brought into each tool. For the zoning press release, P6 wanted a list of each new update to the zoning policy in New York City. She specializes in month-long investigative stories, and incorporating this kind of context would greatly benefit that work. Thus, participants wanted more information pulled into these tools to (1) help them get to interviewing faster and (2) have a deeper understanding of the topic.

8 DISCUSSION

In the following section we discuss a few areas of future work, including enabling journalists to write their own LLM-prompts to customize angle exploration, helping journalists prioritize angles given their time constraints and the likelihood of an angle actually yielding a story, and applying LLMs to read between the lines of other source material, like case law and academic papers. Finally, we end by discussing the limitations of this work.

8.1 Customizing the LLM angle suggestions

Currently, AngleKindling includes a pre-defined set of angles: *controversies*, *negative outcomes*, and *areas to investigate*, but participants expressed interest in customizing AngleKindling to suggest angles that better aligned with their own and their editor's interests. Our user study provides additional evidence for the need of personalization in computational tools for journalists [20]. P4 stated that he normally likes to write stories about "finance or the economy" and that being able to "push the angles in that direction" would be really useful. One way to help journalists personalize their angles could be to help them write their own LLM prompts. However, there are many challenges novices face when writing LLM prompts, including (1) phrasing the prompt so that it best fulfills the task, (2) providing a diverse set of training examples, and (3) scoping the prompt so that it does not ask for too much in one completion [28] [27]. A first step toward helping journalists customize the LLM-prompts could be having them write their own angles as they read press releases. For example, P4 could record financial-impact angles they thought of as they read, as well as highlight the portion of the press release that inspired each angle. These training samples could then be used as examples for a few-shot prompt, similar to the one shown in Figure 2, which could generate financial angles for new press releases. Past work has shown that with support, novices can write their own prompts to create simple AI-applications [60] [59], but has so far been only studied with UX designers and product managers. Future work can examine what specific challenges professional journalists face when writing their own prompts and how to best support them.

Helping journalists write their own prompts could also help them better understand how the system creates its angle suggestions. While using AngleKindling, P7 and P4 both explained that they might trust the suggestions more if they understood how they were generated. P4 was concerned that the angle suggestions might bias him toward certain types of stories and was worried that by spending time examining suggestions, he might be blinded to other angles he might have come up with on his own. Future work can address if letting journalists write their own prompts and familiarizing themselves with LLMs alleviates or exacerbates these anxieties. Perhaps by writing their own prompts, journalists feel they more thoroughly and naturally explore the space of angles, or alternatively, they might realize the LLM's limitations and trust it less as a source for angles.

8.2 Prioritizing different angles based on journalistic constraints

In addition to helping journalists personalize angles that better match their own or their editor's interests, AngleKindling can also support journalists in prioritizing which of these angles to pursue, based on time-constraints or evidence. As explained in the user study findings, AngleKindling provides angles that lend themselves to different types of stories, including day-of, next-day, and months-later. Instead of organizing angles by their type, such as *controversy* or *negative outcome*, they could be organized by how much time and work they would take to fulfill. For journalists with just a day to produce a story, AngleKindling could prioritize angles that can be fulfilled quickly, like public reactions and summarization pieces, instead of more investigative stories that require a deep dive into past legislation or interviews with experts. Another potential avenue is to prioritize angles that are more likely to pique reader interest; this was a feature that P6, P10, and P11 explicitly mentioned would be really useful in a system like AngleKindling. Even with a custom LLM-prompt producing angles that are more aligned with their reader's interests, AngleKindling could support highlighting the most interesting ones from this set. If AngleKindling was deployed in a newsroom, one interesting direction to take is to use click-through rates for articles to train a classifier that could identify which angles would lead to stories their readers might most be interested in. Thus, one rich area for future research is helping journalists sort the system's generated angles to satisfy these important constraints.

As well as sorting angles by projected reader interest and required effort, another concern participants had was determining which angles would actually lead to interesting stories. While P7 thought the *controversies* presented interesting ideas, she said it would be difficult to choose which ones to conduct follow-up research for in practice. For the zoning press release, she pointed at the *controversy*: "The plan could lead to gentrification", and asked, "Why are you feeding me that angle? Why would I go down that route? That would take a lot of time to verify that route". While she thought that gentrification was a potential outcome of the new zoning policies, she had no conception of how *likely* this was the case or if there was any recent evidence that could support this angle. Meanwhile, the provided historical background was a 2015 news article, which provided evidence that past zoning plans did not include enough affordable housing. However, this information was too old and did not help her assess the new plan. Thus, another line of future work could involve understanding how to best gather initial evidence for angles, so that journalists can quickly see which are most viable. Past work has shown user generated content, such as comments and posts on social media, can be filtered to help journalists find sources and information for their stories [57]. A similar strategy can be used to help provide evidence, such as recent posts or users to interview from social media platforms, for angles. Overall, beyond helping journalists realize the many stories that can be written from a press release, future work can investigate how to help prioritize angles that already have evidence to support them.

8.3 The promising ability of LLMs to read between the lines and potential dangers

Our evaluation of AngleKindling provides initial evidence that LLMs can identify the hidden implications of a source text. These implications were sometimes completely unexpected and appreciated by professional journalists, like "There might be infighting among the various agencies involved" and "The plan could lead to more traffic and congestion in New York City." While we applied LLMs to read between the lines for press releases, they can be applied to many other domains where reporters may ground a story in a specific document, such as law and academia. Case law, for instance, would be an interesting source of documents to assess the capabilities of an LLM for unveiling implications. Each case consists of a lengthy reasoning portion that incorporates the relevant circumstances and facts as well as relevant prior law to explain the court's decision. An LLM can be applied to dissect the court's argument and generate implications on (1) how this reasoning might affect the outcomes of similar disputes and more broadly, (2) how this decision might affect our lives. In addition to case law, LLMs can also be applied to unearth the implications of findings in academic papers. LLMs are already being used to help those without a scientific background better understand papers [4] by summarizing findings. But this can be taken a step further to help the authors of these papers explore the ethical implications of their findings, implications for other fields of research, and implications for our daily lives. These are ripe domains for future work in understanding if and how LLMs can help us unearth the implicit connections within a source text.

While LLMs have great potential for reading between the lines and generating implications from text, there are a few implicit dangers in applying them in this way. LLMs are biased by their training data [34], and this bias could appear in the implications they generate. GPT-3, in particular, is trained on a huge amount of text from the Internet, and the news articles it has seen might affect the kind of angles it generates. With the proliferation of click-bait and fear-mongering articles on the Internet, LLMs could potentially skew their angles toward these less desirable directions. Currently, AngleKindling generates angles focused on controversy and negative outcomes, and while this made sense for positively biased press releases, this might lead fear-mongering when applied to other documents. One way to mitigate this issue is to include positive angles, which would identify unexpected or interesting ways people might benefit from the press release. Moreover, after long-term usage of AngleKindling, the system could provide metrics to help journalists self-reflect on the angles they gravitate toward. Perhaps if a journalists selects mostly negative or click-bait-like angles, the system could suggest alternative, positive angles, or at the very least, inform the journalist so that they are aware of their tendencies. In conclusion, another line of future work is to study how to ensure tools like AngleKindling support responsible journalism, by (1) providing angles that encourage deeper journalism as opposed to click-bait articles and (2) encouraging self-reflection.

8.4 Limitations and Future Work

While we carefully designed our study, it is not without limitations. First, our implementation of INJECT does not include all of its features, specifically the ability to search over a large corpus of news articles from a variety of different sources. While we did not include this feature, we do believe that our implementation was enough to compare the two broader strategies of both tools, which for INJECT is providing relevant articles organized by their angle type. That being said, there is the possibility that being able to search over a larger corpus might have improved participants' perceptions of INJECT. Though from the qualitative results we don't think this is the case; participants preferred AngleKindling because it was more "proactive" and provided concrete ideas as opposed to only news articles.

The next set of limitations pertains to our participants. We recruited professional journalists across the United States but had them all come up with angles for press releases from the New York City Mayor's office. This means that many participants lacked the additional context about New York, its prominent politicians, and its history when coming up with angles. However, this is not an unrealistic scenario, as many journalists are plunged into a new area's politics and history when they move or have recently started their career. Future work can examine how useful tools like AngleKindling are for journalists reading press releases that are well within their beat and expertise. Finally, we also only included journalists from the United States, whereas journalists in other countries might have different opinions on what kind of angles they value and how they prefer to be supported when they come up with story ideas.

Next, there are a few ways we can improve AngleKindling's implementation. First, we used GPT-3 to extract the press release's main points, and while this worked well enough for proof of concept, we could likely improve this summary by utilizing a model with a greater input length, like a long-document summarization model [31]. Furthermore, to generate the controversies, we naively split the press release into sections consisting of the maximum number of paragraphs we could fit into GPT-3's input length. However, a key detail of the press release might have been split apart by this naive sectioning, preventing GPT-3 from generating a compelling angle for that detail. In the future, we could include better semantic segmentation to ensure that each section captures a complete, semantic portion of the press release.

Including improving the segmentation of the press release, we could also experiment with different prompts to improve AngleKindling's angles. We could experiment with zero-shot prompts with different phrasings, few-shot prompts with different example sets or more examples from the user study, and learned-prompts through prompt-tuning [32]. To determine the best performing prompt, we could acquire press releases across a variety of domains, including finance, education, entertainment, transportation, etc. and have external raters evaluate the generated angles of these prompts. By doing so, we could (1) improve AngleKindling's angles by incorporating the best performing prompt and (2) have an understanding of how often these prompts can produce good angles and which domain they work best in.

Finally, LLMs are fundamentally limited by their training data. GPT-3 might not function equally well across beats of science, politics and local news. Perhaps LLMs can be fine-tuned or at least have their prompts tuned [32] to support different beats. Next, publicly available LLMs are often not up to date with the latest news. At the time we used it, GPT-3 only had world knowledge up to 2021, limiting its ability to generate angles on more recent events. LLMs also reflect the bias in their training data [34] and future work can elucidate if and how this bias bleeds into the story brainstorming process.

9 CONCLUSION

Informed by a three-month long co-design, we created AngleKindling, an interactive web tool which employs an LLM to help journalists come up with angles for a press release. We conducted a within-subjects study with 12 professional journalists, comparing AngleKindling to a very relevant and recent creativity support tool for journalists, INJECT. We found that AngleKindling was perceived to be significantly more helpful for coming up with ideas, with significantly less mental demand. This was primarily due to AngleKindling (1) helping journalists recognize angles they had not considered, (2) providing angles that were useful for multiple types of stories, (3) helping journalists quickly and deeply engage with the press release, and (4) providing contextualized historical context. Future work can explore how creating their own LLM-prompts might help journalists customize angle exploration and affect their trust of the system, how we might best help journalists recognize the most viable angles within their time-limit, and how LLMs can be used to read between the lines of other source material, like case law and academic papers.

REFERENCES

[1] Tariq Alhindi, Savvas Petridis, and Smaranda Muresan. 2018. Where is Your Evidence: Improving Fact-checking by Justification Modeling. In *Proceedings of the First Workshop on Fact Extraction and VERification (FEVER)*. Association for Computational Linguistics, Brussels, Belgium, 85–90. https://doi.org/10.18653/v1/W18-5513

[2] Salvatore Andolina, Khalil Klouche, Diogo Cabral, Tuukka Ruotsalo, and Giulio Jacucci. 2015. InspirationWall: Supporting Idea Generation Through Automatic Information Exploration. In *Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition* (Glasgow, United Kingdom) *(C&C '15)*. Association for Computing Machinery, New York, NY, USA, 103–106. https://doi.org/10.1145/2757226.2757252

[3] Salvatore Andolina, Hendrik Schneider, Joel Chan, Khalil Klouche, Giulio Jacucci, and Steven Dow. 2017. Crowdboard: Augmenting In-Person Idea Generation with Real-Time Crowds. In *Proceedings of the 2017 ACM SIGCHI Conference on Creativity and Cognition* (Singapore, Singapore) *(C&C '17)*. Association for Computing Machinery, New York, NY, USA, 106–118. https://doi.org/10.1145/3059454.3059477

[4] Tal August, Lucy Lu Wang, Jonathan Bragg, Marti A. Hearst, Andrew Head, and Kyle Lo. 2022. Paper Plain: Making Medical Research Papers Approachable to Healthcare Consumers with Natural Language Processing. https://doi.org/10.48550/ARXIV.2203.00130

[5] Suyun Sandra Bae, Oh-Hyun Kwon, Senthil Chandrasegaran, and Kwan-Liu Ma. 2020. *Spinneret: Aiding Creative Ideation through Non-Obvious Concept Associations*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376746

[6] Frank Bentley, Katie Quehl, Jordan Wirfs-Brock, and Melissa Bica. 2019. Understanding Online News Behaviors. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–11. https://doi.org/10.1145/3290605.3300820

[7] Rishi Bommasani, Drew A. Hudson, Ehsan Adeli, Russ Altman, Simran Arora, Sydney von Arx, Michael S. Bernstein, Jeannette Bohg, Antoine Bosselut, Emma Brunskill, Erik Brynjolfsson, Shyamal Buch, Dallas Card, Rodrigo Castellon, Niladri Chatterji, Annie Chen, Kathleen Creel, Jared Quincy Davis, Dora Demszky, Chris Donahue, Moussa Doumbouya, Esin Durmus, Stefano Ermon, John Etchemendy, Kawin Ethayarajh, Li Fei-Fei, Chelsea Finn, Trevor Gale, Lauren Gillespie, Karan Goel, Noah Goodman, Shelby Grossman, Neel Guha, Tatsunori Hashimoto, Peter Henderson, John Hewitt, Daniel E. Ho, Jenny Hong, Kyle Hsu, Jing Huang, Thomas Icard, Saahil Jain, Dan Jurafsky, Pratyusha Kalluri, Siddharth Karamcheti, Geoff Keeling, Fereshte Khani, Omar Khattab, Pang Wei Koh, Mark Krass, Ranjay Krishna, Rohith Kuditipudi, Ananya Kumar, Faisal Ladhak, Mina Lee, Tony Lee, Jure Leskovec, Isabelle Levent, Xiang Lisa Li, Xuechen Li, Tengyu Ma, Ali Malik, Christopher D. Manning, Suvir Mirchandani, Eric Mitchell, Zanele Munyikwa, Suraj Nair, Avanika Narayan, Deepak Narayanan, Ben Newman, Allen Nie, Juan Carlos Niebles, Hamed Nilforoshan, Julian Nyarko, Giray Ogut, Laurel Orr, Isabel Papadimitriou, Joon Sung Park, Chris Piech, Eva Portelance, Christopher Potts, Aditi Raghunathan, Rob Reich, Hongyu Ren, Frieda Rong, Yusuf Roohani, Camilo Ruiz, Jack Ryan, Christopher Ré, Dorsa Sadigh, Shiori

Sagawa, Keshav Santhanam, Andy Shih, Krishnan Srinivasan, Alex Tamkin, Rohan Taori, Armin W. Thomas, Florian Tramèr, Rose E. Wang, William Wang, Bohan Wu, Jiajun Wu, Yuhuai Wu, Sang Michael Xie, Michihiro Yasunaga, Jiaxuan You, Matei Zaharia, Michael Zhang, Tianyi Zhang, Xikun Zhang, Yuhui Zhang, Lucia Zheng, Kaitlyn Zhou, and Percy Liang. 2021. On the Opportunities and Risks of Foundation Models. https://doi.org/10.48550/ARXIV.2108.07258

[8] Jelle Boumans. 2018. Subsidizing The News? *Journalism Studies* 19, 15 (2018), 2264–2282. https://doi.org/10.1080/1461670X.2017.1338154

[9] 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. *IEEE Transactions on Visualization and Computer Graphics* 20, 12 (2014), 2271–2280. https://doi.org/10.1109/TVCG.2014.2346431

[10] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel Ziegler, Jeffrey Wu, Clemens Winter, Chris Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. 2020. Language Models are Few-Shot Learners. In *Advances in Neural Information Processing Systems*, H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin (Eds.), Vol. 33. Curran Associates, Inc., 1877–1901. https://proceedings.neurips.cc/paper/2020/file/1457c0d6bfcb4967418bfb8ac142f64a-Paper.pdf

[11] Joel Chan, Steven Dang, and Steven P. Dow. 2016. IdeaGens: Enabling Expert Facilitation of Crowd Brainstorming. In *Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion* (San Francisco, California, USA) *(CSCW '16 Companion)*. Association for Computing Machinery, New York, NY, USA, 13–16. https://doi.org/10.1145/2818052.2874313

[12] Joel Chan, Pao Siangliulue, Denisa Qori McDonald, Ruixue Liu, Reza Moradinezhad, Safa Aman, Erin T. Solovey, Krzysztof Z. Gajos, and Steven P. Dow. 2017. Semantically Far Inspirations Considered Harmful? Accounting for Cognitive States in Collaborative Ideation. In *Proceedings of the 2017 ACM SIGCHI Conference on Creativity and Cognition* (Singapore, Singapore) *(C&C '17)*. Association for Computing Machinery, New York, NY, USA, 93–105. https://doi.org/10.1145/3059454.3059455

[13] Minsuk Chang, Leonore V. Guillain, Hyeungshik Jung, Vivian M. Hare, Juho Kim, and Maneesh Agrawala. 2018. RecipeScape: An Interactive Tool for Analyzing Cooking Instructions at Scale. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3173574.3174025

[14] John Joon Young Chung, Wooseok Kim, Kang Min Yoo, Hwaran Lee, Eytan Adar, and Minsuk Chang. 2022. TaleBrush: Visual Sketching of Story Generation with Pretrained Language Models. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI EA '22)*. Association for Computing Machinery, New York, NY, USA, Article 172, 4 pages. https://doi.org/10.1145/3491101.3519873

[15] Sarah Cohen, James T. Hamilton, and Fred Turner. 2011. Computational Journalism. *Commun. ACM* 54, 10 (oct 2011), 66–71. https://doi.org/10.1145/2001269.2001288

[16] Nicholas Diakopoulos. 2020. Computational News Discovery: Towards Design Considerations for Editorial Orientation Algorithms in Journalism. *Digital Journalism* 8, 7 (2020), 945–967. https://doi.org/10.1080/21670811.2020.1736946

[17] Nicholas Diakopoulos, Munmun De Choudhury, and Mor Naaman. 2012. Finding and Assessing Social Media Information Sources in the Context of Journalism. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 2451–2460. https://doi.org/10.1145/2207676.2208409

[18] Nicholas Diakopoulos, Sergio Goldenberg, and Irfan Essa. 2009. Videolyzer: Quality Analysis of Online Informational Video for Bloggers and Journalists. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Boston, MA, USA) *(CHI '09)*. Association for Computing Machinery, New York, NY, USA, 799–808. https://doi.org/10.1145/1518701.1518824

[19] Nicholas Diakopoulos, Mor Naaman, and Funda Kivran-Swaine. 2010. Diamonds in the rough: Social media visual analytics for journalistic inquiry. In *2010 IEEE Symposium on Visual Analytics Science and Technology*. IEEE, 115–122. https://doi.org/10.1109/VAST.2010.5652922

[20] Nicholas Diakopoulos, Daniel Trielli, and Grace Lee. 2021. Towards Understanding and Supporting Journalistic Practices Using Semi-Automated News Discovery Tools. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, Article 406 (oct 2021), 30 pages. https://doi.org/10.1145/3479550

[21] Robert M Entman. 1993. Framing: Towards clarification of a fractured paradigm. *McQuail's reader in mass communication theory* 390 (1993), 397.

[22] Martin Flintham, Christian Karner, Khaled Bachour, Helen Creswick, Neha Gupta, and Stuart Moran. 2018. Falling for Fake News: Investigating the Consumption of News via Social Media. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–10. https://doi.org/10.1145/3173574.3173950

[23] Katy Ilonka Gero and Lydia B. Chilton. 2019. Metaphoria: An Algorithmic Companion for Metaphor Creation. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300526

[24] Katy Ilonka Gero, Vivian Liu, and Lydia Chilton. 2022. Sparks: Inspiration for Science Writing Using Language Models. In *Designing Interactive Systems Conference* (Virtual Event, Australia) *(DIS '22)*. Association for Computing Machinery, New York, NY, USA, 1002–1019. https://doi.org/10.1145/3532106.3533533

[25] Naeemul Hassan, Fatma Arslan, Chengkai Li, and Mark Tremayne. 2017. Toward automated fact-checking: Detecting check-worthy factual claims by claimbuster. In *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining*. 1803–1812.

[26] Gaoping Huang and Alexander J. Quinn. 2017. BlueSky: Crowd-Powered Uniform Sampling of Idea Spaces. In *Proceedings of the 2017 ACM SIGCHI Conference on Creativity and Cognition* (Singapore, Singapore) *(C&C '17)*. Association for Computing Machinery, New York, NY, USA, 119–130. https://doi.org/10.1145/3059454.3059481

[27] Ellen Jiang, Kristen Olson, Edwin Toh, Alejandra Molina, Aaron Donsbach, Michael Terry, and Carrie J Cai. 2022. PromptMaker: Prompt-Based Prototyping with Large Language Models. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI EA '22)*. Association for Computing Machinery, New York, NY, USA, Article 35, 8 pages. https://doi.org/10.1145/3491101.3503564

[28] Ellen Jiang, Edwin Toh, Alejandra Molina, Kristen Olson, Claire Kayacik, Aaron Donsbach, Carrie J Cai, and Michael Terry. 2022. Discovering the Syntax and Strategies of Natural Language Programming with Generative Language Models. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 386, 19 pages. https://doi.org/10.1145/3491102.3501870

[29] Youwen Kang, Zhida Sun, Sitong Wang, Zeyu Huang, Ziming Wu, and Xiaojuan Ma. 2021. MetaMap: Supporting Visual Metaphor Ideation through Multi-Dimensional Example-Based Exploration. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 427, 15 pages. https://doi.org/10.1145/3411764.3445325

[30] Yui Kita and Jun Rekimoto. 2018. V8 Storming: How Far Should Two Ideas Be?. In *Proceedings of the 9th Augmented Human International Conference* (Seoul, Republic of Korea) *(AH '18)*. Association for Computing Machinery, New York, NY, USA, Article 14, 8 pages. https://doi.org/10.1145/3174910.3174937

[31] Huan Yee Koh, Jiaxin Ju, Ming Liu, and Shirui Pan. 2022. An Empirical Survey on Long Document Summarization: Datasets, Models and Metrics. *ACM Comput. Surv.* (jun 2022). https://doi.org/10.1145/3545176 Just Accepted.

[32] Brian Lester, Rami Al-Rfou, and Noah Constant. 2021. The Power of Scale for Parameter-Efficient Prompt Tuning. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Online and Punta Cana, Dominican Republic, 3045–3059. https://doi.org/10.18653/v1/2021.emnlp-main.243

[33] Hang Li. 2022. Language Models: Past, Present, and Future. *Commun. ACM* 65, 7 (jun 2022), 56–63. https://doi.org/10.1145/3490443

[34] Paul Pu Liang, Chiyu Wu, Louis-Philippe Morency, and Ruslan Salakhutdinov. 2021. Towards Understanding and Mitigating Social Biases in Language Models. In *Proceedings of the 38th International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 139)*, Marina Meila and Tong Zhang (Eds.). PMLR, 6565–6576. https://proceedings.mlr.press/v139/liang21a.html

[35] Vivian Liu and Lydia B Chilton. 2022. Design Guidelines for Prompt Engineering Text-to-Image Generative Models. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 384, 23 pages. https://doi.org/10.1145/3491102.3501825

[36] Vivian Liu, Han Qiao, and Lydia Chilton. 2022. Opal: Multimodal Image Generation for News Illustration. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology* (Bend, OR, USA) *(UIST '22)*. Association for Computing Machinery, New York, NY, USA, Article 73, 17 pages. https://doi.org/10.1145/3526113.3545621

[37] Xiaomo Liu, Armineh Nourbakhsh, Quanzhi Li, Sameena Shah, Robert Martin, and John Duprey. 2017. Reuters tracer: Toward automated news production using large scale social media data. In *2017 IEEE International Conference on Big Data (Big Data)*. 1483–1493. https://doi.org/10.1109/BigData.2017.8258082

[38] Ryan Louie, Andy Coenen, Cheng Zhi Huang, Michael Terry, and Carrie J. Cai. 2020. *Novice-AI Music Co-Creation via AI-Steering Tools for Deep Generative Models*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376739

[39] Neil Maiden, Konstantinos Zachos, Amanda Brown, George Brock, Lars Nyre, Aleksander Nygård Tonheim, Dimitris Apsotolou, and Jeremy Evans. 2018. Making the News: Digital Creativity Support for Journalists. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA,

1–11. https://doi.org/10.1145/3173574.3174049
[40] Adam Marcus, Michael S. Bernstein, Osama Badar, David R. Karger, Samuel Madden, and Robert C. Miller. 2011. Twitinfo: Aggregating and Visualizing Microblogs for Event Exploration. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 227–236. https://doi.org/10.1145/1978942.1978975
[41] Changhoon Oh, Jinhan Choi, Sungwoo Lee, SoHyun Park, Daeryong Kim, Jungwoo Song, Dongwhan Kim, Joonhwan Lee, and Bongwon Suh. 2020. *Understanding User Perception of Automated News Generation System.* Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376811
[42] Henrik Örnebring. 2010. Sourcing the News: Key Issues in Journalism-an Innovative Study of the Israeli Press. *Journalism and Mass Communication Quarterly* 87, 3/4 (2010), 682.
[43] Hiroyuki Osone, Jun-Li Lu, and Yoichi Ochiai. 2021. BunCho: AI Supported Story Co-Creation via Unsupervised Multitask Learning to Increase Writers' Creativity in Japanese. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI EA '21)*. Association for Computing Machinery, New York, NY, USA, Article 19, 10 pages. https://doi.org/10.1145/3411763.3450391
[44] Savvas Petridis, Hijung Valentina Shin, and Lydia B Chilton. 2021. SymbolFinder: Brainstorming Diverse Symbols Using Local Semantic Networks. In *The 34th Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) *(UIST '21)*. Association for Computing Machinery, New York, NY, USA, 385–399. https://doi.org/10.1145/3472749.3474757
[45] Zvi Reich. 2006. THE PROCESS MODEL OF NEWS INITIATIVE. *Journalism Studies* 7, 4 (2006), 497–514. https://doi.org/10.1080/14616700600757928
[46] Nils Reimers and Iryna Gurevych. 2019. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*. Association for Computational Linguistics, Hong Kong, China, 3982–3992. https://doi.org/10.18653/v1/D19-1410
[47] Hanieh Shakeri, Carman Neustaedter, and Steve DiPaola. 2021. SAGA: Collaborative Storytelling with GPT-3. In *Companion Publication of the 2021 Conference on Computer Supported Cooperative Work and Social Computing* (Virtual Event, USA) *(CSCW '21)*. Association for Computing Machinery, New York, NY, USA, 163–166. https://doi.org/10.1145/3462204.3481771
[48] Merryn Sherwood, Timothy Marjoribanks, and Matthew Nicholson. 2019. Public Relations and Journalism. https://doi.org/10.1093/acrefore/9780190228613.013.866
[49] Yang Shi, Yang Wang, Ye Qi, John Chen, Xiaoyao Xu, and Kwan-Liu Ma. 2017. IdeaWall: Improving Creative Collaboration through Combinatorial Visual Stimuli. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (Portland, Oregon, USA) *(CSCW '17)*. Association for Computing Machinery, New York, NY, USA, 594–603. https://doi.org/10.1145/2998181.2998208
[50] Pao Siangliulue, Kenneth C. Arnold, Krzysztof Z. Gajos, and Steven P. Dow. 2015. Toward Collaborative Ideation at Scale: Leveraging Ideas from Others to Generate More Creative and Diverse Ideas. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing* (Vancouver, BC, Canada) *(CSCW '15)*. Association for Computing Machinery, New York, NY, USA, 937–945. https://doi.org/10.1145/2675133.2675239
[51] Pao Siangliulue, Joel Chan, Steven P. Dow, and Krzysztof Z. Gajos. 2016. IdeaHound: Improving Large-Scale Collaborative Ideation with Crowd-Powered Real-Time Semantic Modeling. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (Tokyo, Japan) *(UIST '16)*. Association for Computing Machinery, New York, NY, USA, 609–624. https://doi.org/10.1145/2984511.2984578
[52] C. Estelle Smith, Eduardo Nevarez, and Haiyi Zhu. 2020. *Disseminating Research News in HCI: Perceived Hazards, How-To's, and Opportunities for Innovation.* Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376744
[53] John Stasko, Carsten Görg, and Zhicheng Liu. 2008. Jigsaw: supporting investigative analysis through interactive visualization. *Information visualization* 7, 2 (2008), 118–132. https://doi.org/10.1057/palgrave.ivs.9500180
[54] Neil Thurman. 2019. Computational journalism. In *The handbook of journalism studies*. Routledge, 180–195.
[55] Daniel Trielli and Nicholas Diakopoulos. 2019. Search as News Curator: The Role of Google in Shaping Attention to News Information. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3290605.3300683
[56] Hao-Chuan Wang, Dan Cosley, and Susan R. Fussell. 2010. Idea Expander: Supporting Group Brainstorming with Conversationally Triggered Visual Thinking Stimuli. In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work* (Savannah, Georgia, USA) *(CSCW '10)*. Association for Computing Machinery, New York, NY, USA, 103–106. https://doi.org/10.1145/1718918.1718938
[57] 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* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 447, 18 pages. https://doi.org/10.1145/3411764.3445266
[58] David H Weaver, Lars Willnat, and G Cleveland Wilhoit. 2019. The American journalist in the digital age: Another look at US news people. *Journalism & Mass Communication Quarterly* 96, 1 (2019), 101–130. https://doi.org/10.1177/1077699018778242
[59] Tongshuang Wu, Ellen Jiang, Aaron Donsbach, Jeff Gray, Alejandra Molina, Michael Terry, and Carrie J Cai. 2022. PromptChainer: Chaining Large Language Model Prompts through Visual Programming. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI EA '22)*. Association for Computing Machinery, New York, NY, USA, Article 359, 10 pages. https://doi.org/10.1145/3491101.3519729
[60] Tongshuang Wu, Michael Terry, and Carrie Jun Cai. 2022. AI Chains: Transparent and Controllable Human-AI Interaction by Chaining Large Language Model Prompts. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 385, 22 pages. https://doi.org/10.1145/3491102.3517582
[61] Chaolun Xia, Raz Schwartz, Ke Xie, Adam Krebs, Andrew Langdon, Jeremy Ting, and Mor Naaman. 2014. CityBeat: Real-Time Social Media Visualization of Hyper-Local City Data. In *Proceedings of the 23rd International Conference on World Wide Web* (Seoul, Korea) *(WWW '14 Companion)*. Association for Computing Machinery, New York, NY, USA, 167–170. https://doi.org/10.1145/2567948.2577020
[62] Xiaotong (Tone) Xu, Rosaleen Xiong, Boyang Wang, David Min, and Steven P. Dow. 2021. IdeateRelate: An Examples Gallery That Helps Creators Explore Ideas in Relation to Their Own. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, Article 352 (oct 2021), 18 pages. https://doi.org/10.1145/3479496
[63] Lixiu Yu and Jeffrey V. Nickerson. 2011. Cooks or Cobblers? Crowd Creativity through Combination. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 1393–1402. https://doi.org/10.1145/1978942.1979147
[64] 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* (Helsinki, Finland) *(IUI '22)*. Association for Computing Machinery, New York, NY, USA, 841–852. https://doi.org/10.1145/3490099.3511105