"Helping Me Versus Doing It for Me": Designing for Agency in LLM-Infused Writing Tools for Science Journalism

Sachita Nishal*
Northwestern University
Evanston, Illinois, USA
nishal@u.northwestern.edu

Mina Lee*
University of Chicago
Chicago, Illinois, USA
mnlee@uchicago.edu

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

Jennifer Wortman Vaughan
Microsoft Research
New York, New York, USA
jenn@microsoft.com

Abstract

Journalists rely on their *agency*—the ability to exercise independent judgment in alignment with their values—to fulfill their democratic social role. In this study, we investigate how LLM-infused writing tools reshape journalists' agency in editorial decision making. In interviews with 20 science journalists, we presented four hypothetical LLM-infused writing tools representing a range of possible design space configurations. We find that journalists are selectively willing to cede control: they view AI that gathers information or offers feedback as supporting their efficiency by automating execution while leaving decision making intact. In contrast, they see AI that generates core ideas or drafts as a threat to their autonomy, skill development, self-fulfillment, and professional relationships. This sensitivity extends to seemingly automatable tasks such as manipulating writing voice with AI, which are seen as reducing opportunities for reflection and critical thinking. We discuss the implications of these findings for design that preserves journalistic agency in the moment, and over the long term.

CCS Concepts

• **Human-centered computing** → **Human computer interaction (HCI)**; **Collaborative and social computing**.

Keywords

agency, value-sensitive design, journalism, science communication, large language models (LLMs)

*Work done while at Microsoft Research.

1 Introduction

Journalists help sustain informed public discourse by delivering reporting that is timely and relevant. Doing this hinges on journalists' *agency*, their ability to make independent editorial decisions while ensuring that outcomes align with professional values like transparency, accuracy, and public welfare [36]. Journalists exercise their agency within the context of rich sociotechnical systems, frequently using computational tools to support their work [35, 48]. However, the increasing adoption of generative AI in journalism—from information gathering to data analysis to writing headlines and summaries [16]—changes how journalists make and execute decisions and creates risks of value erosion.

Consider LLM-based headline generation, where an editor may use a tool to generate multiple possible headlines, then select and iterate on a subset. This configuration allows editors to explore multiple options quickly, shifting control over initial text generation to the AI system. This may enhance efficiency for breaking news, but if the tool consistently favors engagement-optimized headlines, the editor's selection process could subtly shift over time, perhaps sacrificing accuracy to prioritize audience metrics. This is not a hypothetical concern: practicing journalists worry that generative AI use may reconfigure their practice in ways that divert from their values and diminish their critical thinking and autonomy [12, 55, 59, 63]. Empirical evidence from different knowledge work contexts also lends some credence to these worries: generative AI can cause users to fixate on tool outputs rather than explore diverse ideas [60], shape writing decisions even when its suggestions conflict with users' values [31], and reduce users' critical engagement with tasks [38].

In this study, we examine how journalists' professional values might shape their desired control over LLM-infused tools for writing, i.e., how journalists might negotiate and uphold their *agency* in these sociotechnical configurations. We adopt a value-sensitive design (VSD) approach [19], which centers human values in the design process. This approach directs our investigation toward understanding how people might prioritize values in context, and navigate trade-offs when technological affordances conflict with those values.

We center freelance science journalists who must balance technical accuracy with engaging storytelling, while navigating the unstable economics and relationship-driven nature of freelance work [10]. We investigate these dynamics through the task of pitch writing, where journalists submit short proposals that outline a

story idea and reporting plan to editors. Pitching encapsulates broader value tensions in journalism: it requires creativity within structural constraints and continual negotiation between developing a personal voice and fulfilling editorial expectations. While concerns in pitching and freelancing may not always generalize to other contexts, e.g., journalists might be more willing to experiment with AI-generated phrasing in a short pitch than in a published story with lasting public impact, and freelancers may worry about forging new relationships more so than staff reporters, pitching nonetheless offers a representative window into how journalists weigh their professional values and sense of control when engaging with LLM-infused writing tools.

Our notion of *agency* in sociotechnical systems is grounded in prior work in HCI [5]. We conceptualize agency along two dimensions: *causal agency*, i.e, *control* over decision making and execution with technology, and *identity-based agency*, which captures alignment of work with professional *values* and the shaping of said values over time. Our research questions are:

- **RQ1**: What values underpin journalists' agency during pitch writing?
- **RQ2**: What control are journalists willing to cede to generative AI, and what must they retain, in their pursuit of these values?

To address these questions, we conducted semi-structured interviews with 20 professional science journalists, presenting four hypothetical LLM-infused writing tools that varied in terms of who initiated the task (tool or user), the scope of AI assistance (information gathering, ideation, drafting, feedback), and the degree to which the system could be configured. We drew on design workbooks [21] and speed-dating [68] methods to surface participants' latent needs and explore the boundaries of their desired agency. These approaches are well-suited to emerging technologies with vast design spaces, since they enable exploration of different sociotechnical configurations and user responses.

Our findings show that journalists set clear boundaries around AI assistance in pitch writing, guided by professional values spanning craft concerns (e.g., autonomy, distinctiveness, skill development) and contextual pressures (e.g., editorial relationships, ethical standards). Journalists' willingness to cede or retain control broadly depends on whether AI use preserves their cognitive engagement during pitch writing, which can further support skill development and sense of fulfillment. They welcome AI support for executing tasks like information gathering and feedback generation (e.g., through suggesting resources or edits) as long as they retain control over what to include and how to present it. However, more direct, fine-grained control over decision making and execution remains essential during idea generation and drafting (e.g., trying out different framings or words for the lede[1]).

AI support for seemingly low-stakes tasks, such as generating first drafts, is perceived to reduce opportunities for open-ended exploration and critical thinking where journalists decide what a story could be about and why it matters. Economic pressures may prompt compromises, such as using AI-generated first drafts for quicker-turnaround or more routine stories, but journalists still strive to uphold craft values like originality and voice in those contexts. Some contend that long-term AI use could also subtly reshape how they evaluate sources, gauge outlet fit, cultivate voice, and collaborate with peers, which may in turn reduce opportunities to exercise and develop these skills.

These findings also show that causal agency and identity-based agency can be mutually constitutive. Exercising control, even for routine tasks like searching for interviewees or evaluating outlet fit, helps uphold values and shapes identity over time. Our findings further raise three challenges for VSD: individual variation in what journalists find meaningful, systemic pressures that drive value compromises, and the difficulty of evaluating impacts on professional identity through short-term studies. These are grounded in science journalists' pitch writing, but may extend to other domains where professionals work to sustain their craft under tight constraints, e.g., academic writing or creative writing.

Our contributions include:

- An empirical account of how journalists negotiate agency with LLM-infused writing tools, identifying nine professional values that guide willingness to cede causal agency (i.e., decision-making and execution control).
- Delineation of conditions under which LLM-infused writing tools can threaten journalists' agency: when early-stage suggestions constrain exploration and generated text bypasses reflective work, which limits opportunities for developing professional identity.
- Design implications for agency-preserving LLM-infused writing tools that support efficiency through information gathering and feedback while preserving journalists' decision-making control.
- A characterization of how causal and identity dimensions of agency can reinforce each other, which suggests the importance of maintaining journalists' engagement with their practice and community.
- Broader implications for VSD practitioners, including the need for configurability to address individual variation, attention to systemic constraints that shape use, and longitudinal approaches to capture gradual shifts in practice and identity.

2 Related Work

We draw on two bodies of literature to frame our study of how LLM-infused writing tools might reconfigure journalists' agency. First, we examine *science journalism as a value-laden and creative profession* to situate our empirical focus, outlining the values, constraints, and decision-making contexts that define journalists' work. Second, we review work on *agency and values in generative AI-supported work*, which highlights how design choices in such tools can influence users' control over their work and the alignment of outcomes with professional values.

2.1 Science Journalism as a Value-Laden and Creative Profession

Science journalists communicate newsworthy events (e.g., breakthroughs, discoveries, controversies) to the public, along with important context about the people, processes, and social implications of scientific research. This work requires balancing multiple, often

---

[1]A *lede* (sometimes also spelled as *lead*) is the opening of a news story or pitch that is aimed at grabbing a reader's attention.

competing considerations in the process of discovering and developing news stories: domain values like accuracy, independence, and public interest [13, 36]; news values that determine what is "newsworthy" such as novelty and social impact of research [4, 24]; the topical and format-specific expectations of editors and audiences; and journalists' own creative and topical interests [20]. This work is inherently collaborative, because news stories emerge through ongoing interactions between journalists, sources, and editors.

Science journalists thus engage in a process of collaborative and constrained creativity. As Reich [51] puts it, "the mix of formulaic constraints and creative freedom of journalists places them in a situation comparable to that of jazz musicians […]: the song may be predetermined, but players are free to improvise within a given framework." Creativity permeates news production, from ideating and framing stories, to gathering and verifying information, to experimenting with writing, publication formats, and audience engagement [14, 43]. This constrained creative work increasingly occurs in a freelance context [2, 10]. Freelancers pitch story ideas to editors and independently navigate research and reporting, with minimal institutional support and no guarantees of being commissioned for their extensive time investment in this process. While this creates an "illusion of freedom"—more autonomy over story choice and workflow—structural constraints, such as economic dependence and editor approval for future work, can limit agency in practice [44].

Pitch writing is a particularly revealing site for studying journalists' values and agency. A pitch is a short proposal that freelance journalists write to convince editors to commission their story ideas [3]. While pitches follow well-defined structural conventions (e.g., including a lede and reporting plan[2]), they demand significant creative work to convince editors that a story is newsworthy and that the particular journalist is the right candidate to tell it. This involves developing fresh angles, showcasing writing voice, and calibrating writing style to match an editor's expectations [58]. This combination of structure and creative freedom makes pitch writing a microcosm of the broader engagement with values, creativity, and agency in journalism.

In this study, we rely on pitch writing as a site to understand how LLM-infused writing tools might reconfigure journalists' agency. Following Bennett et al. [5], we conceptualize agency through its *causal* aspects (decision-making and execution control with technology) and its *identity-based* aspects (alignment with professional values and shaping of professional identity). In journalism, causal agency manifests as control over developing, framing, and presenting story ideas. Grounding this work in value-sensitive design, we attend to how design choices can support or subtly undermine these values, and how practitioners might respond to value trade-offs in design. The stakes are particularly high given how previous technologies adopted in journalism have created new forms of constraint alongside their affordances, from analytics tools that reshape editorial judgment to platform dependencies that make newsrooms vulnerable to algorithmic changes [11]. By focusing on LLMs in pitch writing, we surface where design can reinforce journalists' agency in balancing structural constraints with creative freedom.

---

[2]A reporting plan might detail how long the story will be, what format it will assume, who the writer plans to interview, any graphics or visualizations they plan to include, and so on.

2.2 Agency and Values in Generative AI-supported Work

Generative AI tools, by design, reshape how professionals make and execute decisions in knowledge work. Rather than engaging directly in creation and synthesis, professionals increasingly find themselves orchestrating AI systems [50]. Knowledge workers report that AI use for writing documents and code transforms their processes from emphasizing information gathering and synthesis, toward selecting and integrating AI-generated outputs [38]. This represents what recent work describes as "relocated agency," where professionals cede aspects of control over their work to AI systems through cycles of prompting, iterating, and pushing back on these systems' outputs [28]. Within journalism workflows, these shifts in control take the form of evaluating AI suggestions of news angles and metaphors; guiding scaffolded AI generation of video content; working through AI-generated proofreading suggestions; and so on [1, 34, 49, 61].

However, these process changes can subtly shift professionals away from their intended goals and professional values. Recent work demonstrates that AI can influence what people write even when suggestions conflict with their own beliefs [31], while knowledge workers using AI with greater confidence show reduced critical thinking [38]. Other studies with writers find that content generation can undermine both the thinking processes and authentic self-expression that writers value [29, 40]. These shifts may occur through subtle mechanisms like fixation on AI outputs rather than exploring diverse original ideas [60], and through what researchers identify as the "intentionality gap," where users struggle to evaluate AI outputs for alignment with their goals and may skip this evaluation entirely [57].

Journalists have expressed concerns about how these changes might affect their professional practice. Studies document worries that generative AI use could diminish critical thinking and independent decision making in editorial work [12, 59, 63]. Recent research with journalists across major news organizations reveals that AI contributes to the "rationalization of news work," with journalists worried about being "implicitly steered away from core values" through AI-driven pressures toward efficiency [55]. The economic precarity of freelance journalism heightens these tensions, as journalists must balance efficiency gains against maintaining editorial control and professional standards [10].

Even when professionals recognize these risks, recent work shows that preserving agency when it is relocated requires continuous, active effort. Writers create explicit barriers to AI use when prioritizing certain values [6], describe putting significant energy into maintaining their authentic expression [23], and must actively resist AI's universalizing tendencies so they can produce meaningful work [28]. Others also find that the ability to steer generative AI systems for writing, e.g., through extensive prompting and configuration, may influence feelings of ownership to some degree [32, 65]. Yet the amount of effort users invest may not translate straightforwardly into preserved agency: recent work found that writers who edited and adapted paragraph-level AI suggestions still reported lower ownership and satisfaction than those who received only sentence-level assistance, even though their writing quality improved [15].

If user effort alone cannot preserve agency, then design must focus on creating conditions where professionals can sustain their

values and control even amid relocated agency in LLM-assisted writing. While specific design choices show promise for preserving user agency—such as cognitive forcing functions that require users to make decisions before seeing AI output [57], designing for greater cognitive engagement through low-fidelity outputs that require significant human editing [60, 66], or recommending sources for writing inspiration rather than generating text [29]—most existing studies focus on general creative writing contexts. Journalism, however, is not just another creative profession: it depends on independence and critical judgment, and economic precarity amplifies the tensions between efficiency and these values [46]. Yet we lack systematic accounts of how generative AI design choices might intersect with these agency dynamics in the context of LLM-infused writing tools.

Our work addresses this gap by applying a two-dimensional agency framework to examine how different LLM-infused tool designs might reconfigure both the causal and identity-based aspects of journalists' agency, grounding the analysis in the concrete case of pitch writing in science journalism. In doing so, we show how design can impact journalists' values and offer design implications for other professions navigating similar value tensions.

3 Methods

Our study investigates how journalists exercise agency during pitch writing, and how this shapes their receptiveness to different LLM-infused writing tools. We adopted a research-through-design approach grounded in value-sensitive design (VSD) [19], drawing on design workbooks [21] and speed-dating methods [68] within semi-structured interviews. VSD guided our design process: we crafted speculative concepts that might surface tensions between different professional values (rooted in prior work [14, 35, 48]). We aimed to understand what participants professed to value, and which trade-offs they would accept or reject in context.

Design workbooks present multiple, provisional, and often ambiguous design concepts to stimulate reflection and speculation, inviting participants to project their own contexts and values onto potential futures from a design space [21]. Speed-dating, a complementary method, involves rapidly exposing participants to several design concepts to elicit feedback on desirability and feasibility, and to surface unmet needs or social tensions [68]. We drew on these approaches to develop four LLM-infused pitch writing concepts, which served as instruments in semi-structured interviews with freelance journalists. This approach is particularly useful for emerging technologies like generative AI, where the design space is vast and evolving, and early-stage user engagement can help designers understand how professional values might shape the acceptance and use of novel tools.

Our approach follows recent HCI work using design workbooks and speed dating to investigate novel technologies in specific contexts [6, 9, 27]. We allowed for slightly longer engagement with each concept than typical speed-dating studies to accommodate the depth of discussion around how these tools might interact with journalists' values, control, and ultimately, agency. The study protocol was approved by our institutional ethics board.

3.1 Developing the Design Concepts

We developed a design workbook containing four hypothetical LLM-infused tools for science journalists to support pitch writing. A pitch is a short, persuasive proposal that freelance journalists submit to editors, combining research, story ideation and framing [3]. In encompassing these activities, pitch writing represents a microcosm of journalism's different value tensions: balancing individual creativity with professional norms, negotiating editorial expectations, and managing efficiency under time constraints. While some priorities may shift in the writing of full stories, the pitch offers a useful lens for examining how journalists navigate agency within new sociotechnical arrangements.

The design team included researchers with experience in building tools for journalists, LLM-infused writing systems, and responsible AI design. Our iterative design process was informed by prior work mapping LLM-infused writing tools across dimensions of variation like tasks, users, technology leveraged, interaction modes, and deployment ecosystems [39]. To ground the design in journalists' practice, we drew on literature characterizing the pitch writing process [3, 17, 25] and examined real-world pitch examples.[3] This helped us understand the key activities involved, such as background research, angle development, framing, and the constraints journalists navigate, such as outlet-specific guidelines, editorial expectations, and time pressures.

In early explorations, we generated a wide range of possible LLM-based features addressing different pitch writing needs, including idea generation, copy editing and style transfer, information sourcing, and templated writing. We sought to vary control across multiple dimensions: control over the writing process (which subtasks to support), control over system features (degree of configurability), control over input modality (natural language vs. structured options), and control over model outputs (semantic qualities like tone vs. mechanical constraints like length).

Through several rounds of brainstorming and mapping, we narrowed our focus to three dimensions that could be systematically varied to create meaningfully different experiences: (1) **Task Initiation**: who initiates AI involvement; (2) **Scope of AI Assistance**: which functionality the AI offers; and (3) **System Configurability**: the degree of control journalists have over AI output. We selected features that might maximize both usefulness and differentiation across concepts, allowing us to probe different value tensions. Guerrilla testing with colleagues and journalists (n=6) over four weeks helped us refine concepts to ensure they remained plausible while spanning this three-dimensional space. The final four concepts were positioned at different points across these dimensions, with each embodying hypotheses about how design decisions might shape journalists' agency and their navigation of professional values.

We created high-fidelity mockups rather than functional prototypes, consistent with prior design workbook methods [6, 9]. Mockups were implemented as interactive Figma prototypes with preset interaction paths, allowing participants to navigate through realistic usage scenarios while keeping the focus on potential outcomes and implications rather than technical implementation details. Our workbook embodied multiplicity, ambiguity, and provisionality [21], presenting a range of system configurations, provoking both

---

[3]Available here: https://www.theopennotebook.com/pitch-database/

Table 1: The four design concepts used in our study, varying in task initiation, scope of AI assistance, and system configurability.

Design Concept	Task Initiation	Scope of AI Assistance	System Configurability
Pitch Assist	User-initiated via open-ended chat interface	Information gathering, ideation, drafting, feedback generation	Open-ended prompting; full text editing; no scaffolding
Pitch Refine	User-initiated via scaffolded UI with subtasks	Information gathering, ideation, drafting	Parameter tuning (e.g., tone, length); structured prompts; modular inputs
Pitch Critic	System-initiated feedback triggered during writing	Feedback generation	Limited control; choose feedback mode; accept/reject inline suggestions
Pitch Suggest	System-initiated automatic drafting based on topic and opportunities	Ideation, drafting	Minimal configuration; preset prompt style; minor revision allowed post-output

interest and resistance, and allowing journalists to reflect on their practices and articulate where they might welcome AI involvement (or not).

3.2 Design Concept Descriptions

Below we describe our four speculative design concepts (Table 1) to explore how AI systems might reconfigure agency across different points in the pitch writing workflow, with references to trace inspiration from prior exemplars. To situate these concepts, recall that we examine agency through users' *causal* (decision-making and execution control) and *identity* (alignment with and shaping of professional values) dimensions, which each concept engages differently.

Pitch Assist offered a user-initiated, open-ended chat interface supporting information gathering, ideation, drafting, and feedback (Figure 1), inspired by multi-modal chat interfaces. In a scenario around pitching a short, environmental science story based on a press release, participants were shown how they could use an LLM to brainstorm angles, headlines, reader concerns, and visualizations, as well as how they could manipulate the style and length of the LLM-generated writing. The design entailed high configurability, allowing journalists to shape prompts, edit freely, and decide when to engage the tool. We envisioned that this concept could support decision-making control well, since users initiate and guide the process, but execution control may be more variable, depending on how users envision relying on system outputs to shape their work.

Pitch Refine featured a scaffolded, user-initiated interface organized into sub-tasks supporting ideation, information gathering, and drafting (Figure 2). The scenario was based on a short tech-focused story derived from a research paper, where participants could select from predefined, task-specific steps [67] like refining a lede, brainstorming a reporting plan, or adjusting voice and tone. The system offered moderate configurability through structured prompts and parameter tuning (e.g., for tone, length) [66], helping users focus their requests while constraining freeform prompting. We envisioned that this design might afford moderate decision-making control, as users chose when and how to engage with each component. We also envisioned moderate execution control, since users could influence tone and direction, but much of the process was guided through predefined interactions.

Pitch Critic offered system-initiated feedback during writing, focused on evaluative tasks like clarity, accuracy, and bias (Figure 3).

The feedback would also be based on domain-specific text like scientific references shared by users or prior news stories at outlets they were pitching. The scenario involved a health-focused story where participants explored system-generated suggestions tailored to these journalistic criteria, with an optional prompt for custom criteria [66]. While users could choose which feedback types to enable and whether to accept suggestions based on provenance [37, 62], the timing and delivery of assistance were controlled by the system. We envisioned that this concept could still support relatively high decision-making and execution control overall: although feedback was system-initiated, users would have to make many active writing decisions to reach those intervention points, and could retain discretion over how to interpret and apply system suggestions.

Pitch Suggest supported idea generation and drafting based on the journalist's writing history, thematic interests, and pitching style (Figure 4), i.e., personalization. The scenario focused on a longer, mental health-focused story, in which participants could share contextual information and select a target outlet, after which the system automatically generated a draft aligned with their profile. System configurability was minimal since users made a few upfront selections [47] and could revise the output, but had little influence during generation. We envisioned that this concept would afford relatively low decision-making and execution control, as the system led the creative process end-to-end, positioning users primarily as reviewers or editors of AI-generated content.

3.3 Participant Recruitment

We recruited 20 professional science journalists who were current or former freelancers through purposive sampling [45]. We focused on freelance science journalists because pitching is more critical for them than newsroom staff. They must prove both their skill and story newsworthiness with each pitch while building long-term editor relationships [58]. We recruited via science journalism listservs (e.g., Science Writers in New York, SoCal Science Writing) and social media platforms (X, LinkedIn, BlueSky). Participants received $75 for 60–80 minute interviews. We screened participants through an eligibility form requiring them to be English language writers. We also asked participants to describe their reporting experiences with links to bylines to screen out bots and inauthentic respondents.

Our participants wrote across various formats (e.g., short news stories, features, profiles) and beats (e.g., tech, health, nature) with
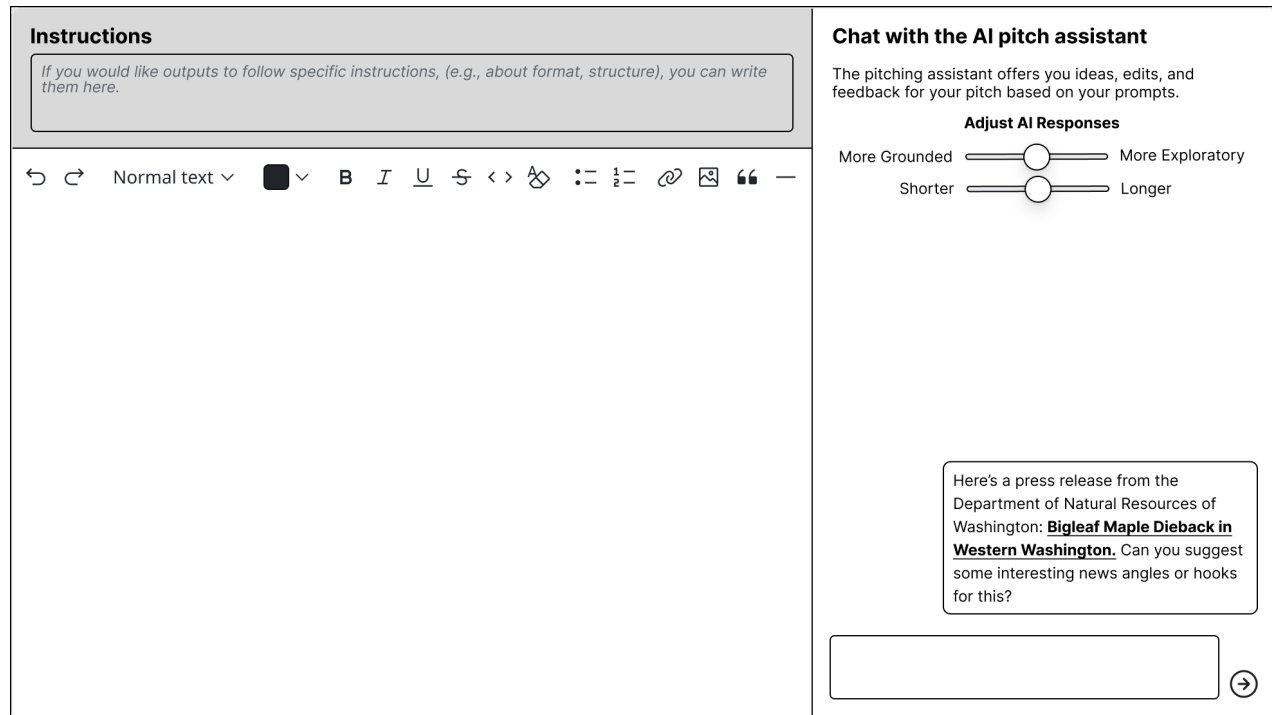
Instructions

If you would like outputs to follow specific instructions, (e.g., about format, structure), you can write them here.

Normal text ⌄ | B *I* U̶ ᵗ̶ ‹ › ⚹ ⋮≡ ₁≡ ⌗ 🖾 ❝ —

Chat with the AI pitch assistant

The pitching assistant offers you ideas, edits, and feedback for your pitch based on your prompts.

Adjust AI Responses

More Grounded ⬤ More Exploratory

Shorter ⬤ Longer

Here's a press release from the Department of Natural Resources of Washington: **Bigleaf Maple Dieback in Western Washington.** Can you suggest some interesting news angles or hooks for this?

Figure 1: Pitch Assist, a chat-based assistant for idea generation, drafting, and revision.

↶ ↷ Normal text ⌄ | B *I* U̶ ᵗ̶ ‹ › ⚹ ⋮≡ ₁≡ ⌗ 🖾 ‹/› ❝ —

Pitch: Don't "Just Google It"! Online searches reinforce misinformation, new study finds

Hi Rachel,

Imagine a world where the more you try to fact-check information online, the more likely you are to believe falsehoods. A **new study** reveals this counter-intuitive reality: using search engines to verify news articles can actually increase belief in misinformation, especially when results lead to low-quality sources.

With the 2024 U.S. elections on the horizon and misinformation concerns at an all-time high, this study challenges the foundations of digital literacy education and fact-checking practices. Across five experiments, the authors find a consistent increase in belief of false news articles among those who searched online to evaluate their truthfulness.

I propose an 800-word news story for MIT Technology Review's website, focusing on the study's implications for how we teach people to check and combat online misinformation. I've conducted preliminary interviews with the study's lead author and will interview a rep from a major search engine company for their perspective on the challenges identified. I can deliver this piece within one week of commissioning.

I'm a technology journalist with a focus on the societal impacts of digital platforms, here are some clips of my work in this area: **clip1**, **clip2**, **clip3**

Refine Your Pitch with AI

Identify different parts of the pitch you want to work on with support from the AI system. The AI system's suggestions are based on any links you provide in your pitch.

[**Lede**] [Plan] [Voice]

Current lede

Imagine a world where the more you try to fact-check information online, the more likely you are to believe falsehoods. A new study reveals this counter-intuitive reality: using search engines to verify news articles can actually increase belief in misinformation, especially when results lead to low-quality sources.

[Choose News Angle ⌄] [Choose Lede Style ⌄] ⊕
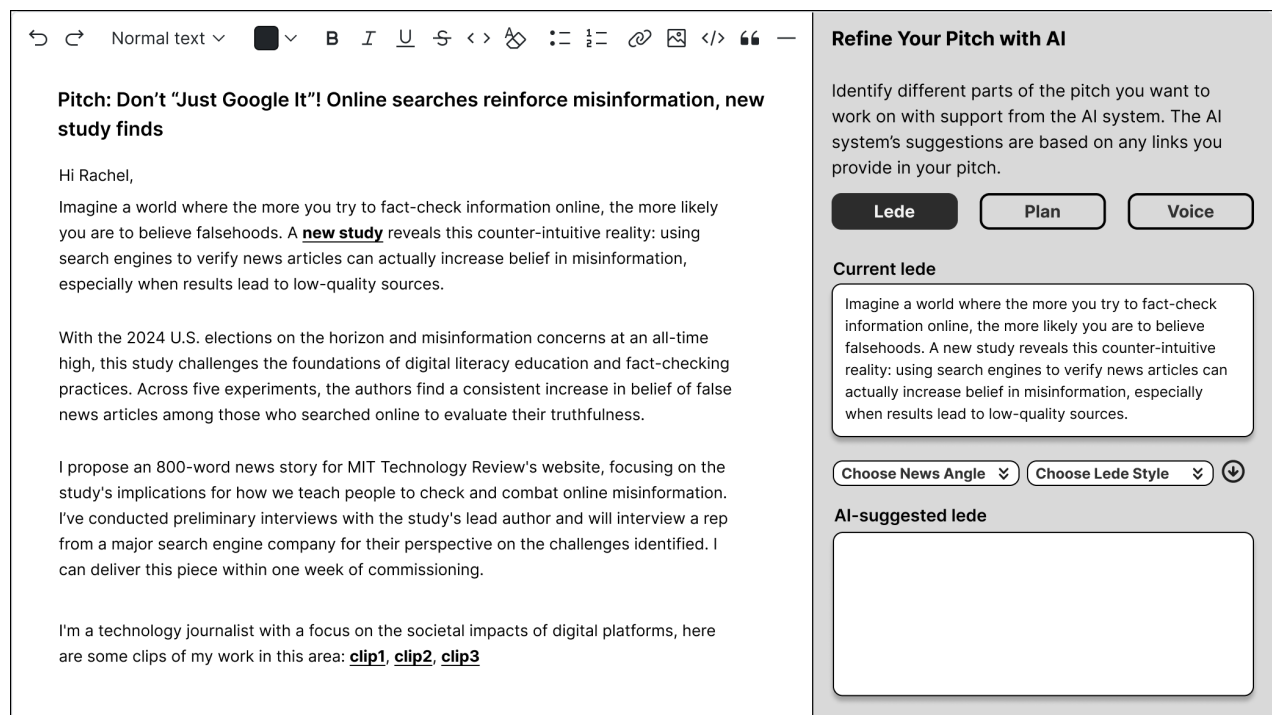
AI-suggested lede

Figure 2: Pitch Refine, a scaffolded interaction for idea generation and revision.

Customize Critique
Choose criteria for which you would like to receive sentence-level critique from the AI system.

☑ **Accuracy**: Based on references provided by you in-text
☑ **Bias**: Excessive sensationalism or negativity toward a topic
☑ **Clarity**: Ambiguous phrases that can impede understanding
☐ *Custom criteria* [*e.g. "potential to add an angle for American outlets"*]

↶ ↷   Normal text ⌄   ▪ ⌄   **B** *I* U̲ S̶ ‹ › 🖉   ⋮≡ ₁≡   🔗 🖼 ‹/› 66 —

Title: Weight Loss Drugs May Be Secret Weapon Against Cancer
Hello,

I'd like to pitch a news story on an unexpected discovery: GLP-1 medications, primarily known for treating obesity and diabetes, may also be a powerful tool in reducing cancer risk.

A new study presented at the recent American Society of Clinical Oncologists (ASCO) meeting has revealed striking results. The research shows that patients in the U.S. taking GLP-1 medications experienced a 39% reduction in risk for obesity-related cancers over time. This dramatic decrease outperforms other weight loss interventions, potentially revolutionizing cancer prevention strategies.

I would interview one of the study's authors and a representative from a pharma company developing these drugs. If needed, I also have contacts at weight loss clinics who could provide perspective on patient experiences.

Editorial Critique

Receive critique on how newsworthy and relevant an editor might find your pitch.

The AI system generates this critique based on pitch guidelines and published news stories of the news outlet you want to write the pitch for.

Choose News Outlet

STAT News ⌄   →

Potential Editorial Feedback

This pitch aligns with STAT's focus on original health and scientific reporting. To strengthen it:

1. Include independent experts for critical analysis and broader context
2. Provide a unique angle to differentiate from other outlets' coverage
3. Explore business implications for GLP-1 drug developers (good for STAT+ readers)
4. Clarify the study's key findings and limitations

This approach would create a well-rounded, STAT-worthy story.

Figure 3: Pitch Critic, a system offering in-text and overall suggestions based on journalistic criteria.

Reporting Record ⊕
This data helps the AI system to suggest pitches most relevant to your interests.

Bookmarks ✕   Writing ✕
Pitches ✕

AI-detected Reporting Interests 🖉
The AI system classifies your interests using these tags. You can delete these tags, or add news ones about your topic or format interests.

Internet ✕   Social Media ✕   Mental Health ✕
Hard News ✕   Explainers ✕

Suggested Call to Pitch

Based on your reporting history and interests, the AI system has suggested this recent call for pitches made on Twitter.

Call from SELF Magazine [Twitter]
- **Topics:** important questions about mental, body, and sexual health
- **Format**: service pitches, e.g., explainers, guides, science
- **Rate**: not mentioned
- **Pitch Guidelines**

Suggested due to your interest in:

Mental Health   Explainers

Suggest Different Call

↶ ↷   Normal text ⌄   ▪ ⌄   **B** *I* U̲ S̶ ‹ › 🖉   ⋮≡ ₁≡   🔗 🖼 ‹/› 66 —

Pitch — Protecting Mental Health on TikTok: A Guide for Young Women

Hi [insert editor name],

I have a story idea that would work well for SELF Magazine. I'm proposing a service piece on safeguarding mental health while using TikTok, **inspired by a new study** revealing that TikTok can negatively affect women's body perception, especially when exposed to "pro-anorexia" content. While social media's impact on mental health isn't new, this study shows the rapid onset of negative effects - within just 10 minutes of use.

The 1000-word article will describe the study's findings about TikTok's impact on women's body image, suggest ways to recognize and avoid harmful content online, and offer tips for mindful social media use. I'll interview the study's lead researchers, an expert from the National Eating Disorder Association, and a social media literacy expert.

I am new to journalism (clips **here** and **here**). I have an M.S. in Psychology from NYU, and I'm committed to destigmatizing mental health issues.

Suggest Different Pitch   **View Past Pitch Suggestions**

Figure 4: Pitch Suggest, a system for idea and draft generation based on journalists' work history.

differing levels of reporting experience, tech expertise, and AI attitudes. Most were North American, with others from Europe, Asia, and Africa. They worked for diverse publications including general interest, science-focused (WIRED, Scientific American, Eos), trade publications (Science, Nature), and local news. Three participants had editing experience and shared insights from evaluating pitches. Appendix A summarizes participant details.

3.4 Interview Protocol

Interviews were conducted via Microsoft Teams. The protocol was developed concurrently with the design workbook through iterative discussions within the research team over the course of four weeks and refined through six pilot interviews. It consisted of three components with distinct goals:

- **Onboarding** (10–15 minutes): The interviewer collected information about the journalists' backgrounds and practices, including their beats, formats, workflows, work goals, challenges, technical expertise, and attitudes toward generative AI tools in their overall workflow.
- **Design Concept Exploration** (50–60 minutes): The interviewer walked participants through each of the four design concepts with a scenario and mock pitch that showcased how its features could help them brainstorm, write, and/or revise a mock pitch[4]; followed by specific questions regarding the design concept and its relationship to their control, values, and practices. Design concepts were presented in random order to mitigate any ordering effects.
- **Debrief** (5–10 minutes): Participants were asked to rank the four concepts by how likely they might be to use them, with explanations.

The complete design workbook and interview protocol are available in the Supplementary Materials.

3.5 Data Analysis

We conducted reflexive thematic analysis using a latent, inductive approach [7, 8] to understand how journalists' responses revealed their professional values and preferred dynamics of control over generative AI systems. The first author conducted initial coding on professionally transcribed interviews. Initial codes were organized into three broad categories reflecting our starting research questions: (1) how journalists conceptualized pitch writing and the values they associated with it; (2) value impacts and tensions that emerged when using AI tools; and (3) design patterns and features participants found desirable or concerning. Memos were created for each code to track emergent patterns.

Initial codes were reviewed by the research team, followed by iterative refinement through clustering, splitting, and merging codes. As we established connections across these initial categories—linking how participants' views of pitching shaped their responses to specific features and the values they invoked—a more direct organizational structure emerged. Rather than separating "value impacts and tensions" as a distinct category, we found it more analytically useful to frame findings around (1) values that underpin journalists'

agency in pitch writing, and (2) boundaries around where journalists would cede or retain control over AI tools in pursuit of certain values. This two-part structure aligned more directly with our theoretical framing of agency while staying grounded in the data.

As an example, early codes about *concerns of voice with AI* initially categorized as a value impact were eventually split and integrated differently: themes about *distinctiveness* as a professional value (Section 4) and specific boundaries about retaining writing control to preserve voice (Section 5.4). Conversely, separate codes about *pitching as unpaid labor* and *pitching as a way to earn* were merged into a broader theme about *efficiency*, as participants emphasized this underlying value when discussing the economic realities of pitching. Design implications that had been coded separately were woven throughout the findings, wherever relevant to specific control boundaries. These themes were refined through regular group meetings over 2–3 months. Throughout the analysis, we remained attentive to different temporalities of agency as emphasized by Bennett et al. [5], i.e., immediate user experience versus longer-term implications.

4 RQ1: Values Underpinning Journalists' Agency in Writing

Our findings show that pitching encompasses a range of activities, from gathering information on newsworthy items and target outlets, to drafting the pitch and iterating based on feedback from editors and peers. Through thematic analysis of our interview data, we surfaced nine values that guide journalists' decisions about what to pitch and how to pitch it, i.e., values that shape their identity-based agency in writing.

These values stem from journalists' treatment of writing as their *craft* (autonomy, writing as thinking, fulfillment, distinctiveness, skill development), as well as journalism's social and commercial *context* (ethical standards, editorial alignment, relationship building, efficiency). While craft and commercial context are often framed as conflicting forces in creative industries [26], our analysis reveals how these values also reinforce each other in practice. In this section we define these different values, situate them within participants' pitching activities, and delineate connections across them of mutual support and conflict as they arise.

Autonomy. Participants expressed a desire for self-direction in their creative process when pitching, such as freedom to choose what newsworthy events to pitch, what news angles or framing to pursue, and how to develop their ideas. This encompasses both autonomy over creative products (the actual material of the pitch and its style) and creative processes (the freedom to explore ideas and write at their own pace without disruption to their natural workflow). As P17 explained, this self-direction is a path to personal fulfillment in work: "You're proposing your own perfect project or a project that's completely self-driven [. . .] it really helps me to connect with the things that I care about in this space and the work that I want to do." Craft values can hence reinforce each other in practice.

Writing as Thinking. Though pitching ostensibly functions to sell stories, participants emphasized it was also an iterative, creative process where they were actively brainstorming stories, seeking out new information, and making connections across seemingly

---

[4]Mock pitches were written by the first author and spanned different topic areas within science and technology to communicate the broad applicability of the design concepts.

disparate ideas. As P7 noted, pitching served as a way to "work out a lot of the kinks in your own thinking." This process of deliberation was deemed valuable regardless of whether the pitch was ultimately accepted. P11 explained: "It can be nice to outline an idea [...] do some of that kind of brainstorming and organization that you would do no matter what." Though some participants preferred to reserve their cognitive efforts for the actual story writing stage, especially for shorter, news story pitches based on press releases that "already have an angle" [P16], most viewed pitching as an essential thinking practice that prepared them for deeper reporting down the line.

Self-fulfillment. Participants strongly valued the intrinsic satisfaction that came from engaging in pitching, finding fulfillment across different stages of the process. Some enjoyed finding untold stories or engaging in background research: "You learn about a new topic or find some exciting research finding" [P2]. Others appreciated the analytical challenge of developing angles: "I like that it's a bit of a puzzle [...] thinking about the different ways that an idea could be framed and which framing is going to make the pitch do the best" [P8]. The creative freedom of pitching also provided satisfaction, as P6 noted: "pitching is more fun than writing [...] the story can be a lot of things that you imagine, and it doesn't have to be pinned down by reality yet." Others enjoyed the mechanics of writing and editing: "you have to be concise and you have to get to the point across [...] it forces you to kind of figure out what is the most important part of the story, and that can be enjoyable" [P9].

Distinctiveness. Through their pitches, participants sought to develop and maintain a specific perspective and writing style while bringing unique or new information to audiences, i.e., distinctiveness of both their voice and ideas. Voice was central to journalists' identity, as P8 described: "Voice is one of the things that you try as a journalist to develop and that you're sort of proud of." Yet this craft value also serves contextual needs through attracting opportunities, as P6 noted: "I know in my work as a writer, working with editors, they are interested in a specific voice that I have." Similarly, participants valued finding stories that "need to be told and haven't been told," [P12], or that presented a new angle on a familiar topic. Distinctiveness thus exemplifies how craft and contextual values reinforce each other, by fulfilling journalists' desire for authentic expression while creating the market differentiation necessary for professional success.

Skill Development. Participants emphasized that pitching offered opportunities to hone their writing craft and deepen subject matter expertise, and so it supported both personal fulfillment and career advancement. As P4 explained: "Pitching is a chance for me to flex my writer-y muscles [...] This is some of the only writing I do that doesn't get edited in at least some capacity. So it's a good way for me to build my skills." Even though writing pitches could be challenging, participants recognized the developmental value, as P12 noted: "Part of writing a pitch helps you understand your own story. [...] It's hard and it's painful. It's like going to the gym. If you don't like exercising, going to the gym is a necessary evil. It's good for your health. But you can't have someone do it for you."

Journalistic Ethics. Participants strove to uphold professional journalism standards such as accuracy, transparency, and lack of bias. They were committed to representing information accurately and providing appropriate context even in the condensed pitch format. As one participant noted: "I don't like simplifying stuff in a way that is misleading" [P3]. However, tensions sometimes arose between writing style and ethical standards, e.g., a phrasing choice in a story title might conflict with absolute accuracy. Participants also emphasized the importance of personal integrity in editor relationships, e.g., through ensuring attribution of their ideas and honest representation of their writing skill, viewing it as essential to long-term professional relationships.

Efficiency. Pitching is the primary mechanism through which freelancers earn money, creating pressure to balance craft values against economic realities. As P13 explained: "My number one goal is to sell an article. I've got to pay the bills." This pressure intensifies because pitching itself is unpaid labor that requires significant time investment, as P16 noted: "It's time-consuming and it involves scanning through a lot of published news and published research to figure out what's really happening [...] for freelance journalists like me, it's a lot of voluntary time" [P16]. The need to tailor pitches for different outlets further compounds this challenge, making efficiency a constant consideration in balancing creative fulfillment with financial necessity.

Editorial Alignment. Successful pitches require alignment with editor preferences and audience needs through finding the right outlet for a story idea, and adapting it to fit publication requirements. This understanding develops through years of experience navigating publication landscapes and reader expectations. As P20 explained: "You're just constantly fishing around [for place to pitch to]. Particularly for the first 10 years or so, I was still getting a handle on what the landscape looks like." This alignment also requires journalists to balance their personal interests with market demands, adapting their voice and focus to meet external expectations while maintaining their authentic perspective.

Relationship Building. Pitches serve as transactions or creative explorations, but also as opportunities to develop professional relationships with editors. Journalists expressed hoping to become established enough that editors and sources would approach them directly with opportunities. As such relationships mature, pitching becomes "kind of more of a collaborative process with editors" [P7], with experienced journalists directing more effort toward managing existing relationships rather than pitching speculatively. Beyond editor relationships, several participants also valued building connections within the broader journalism community through mentorship and peer networks.

5 RQ2: Dynamics of Control with LLM-Infused Writing Tools

Participants' responses to the different design concepts revealed consistent patterns around four types of AI functionality: information gathering (surfacing opportunities, sources, resources for pitching), idea generation, text generation (for different parts of the pitch), and feedback (pitch evaluation and suggestions). Participants readily accepted AI assistance for information gathering (Section 5.1) as long as they maintained control over evaluation and selection. For feedback generation (Section 5.2), they welcomed optional flagging and critiques of their writing based on domain-specific criteria. In contrast, AI support for idea (Section 5.3) and text generation (Section 5.4) was treated more cautiously: features that kick-started thinking were acceptable, but those generating

core ideas or substantial text threatened craft values like autonomy, reflection, voice, and fulfillment.

These preferences were reflected in participant rankings, where Pitch Critic was ranked first by half the participants (n=10), followed by Pitch Suggest by 5 participants, with those ranking Pitch Suggest first explicitly valuing its information-gathering capabilities over its automated drafting features. The remaining 5 participants ranked either Pitch Refine or Pitch Assist first, and 4 of them then ranked Suggest or Critic second (Appendix B). However, individual variation also emerged: some participants rejected AI assistance even for typically accepted tasks (information gathering, feedback) when those activities contributed to their enjoyment or skill development, while others accepted typically rejected features (idea or text generation) when facing contextual pressures like financial constraints, tight deadlines, or fear of rejection.

5.1 Connecting Interests, Demands, and Resources Through Recommendations for Information Gathering

Participants actively desired AI assistance for information gathering during pitching: they were willing to cede execution control over searching for and aggregating resources, but insisted on retaining decision-making control over evaluation and selection of what information to use in their pitch and how. This pattern supported efficiency and editorial alignment, without impinging on participants' sense of autonomy, fulfillment, or originality. Several participants explicitly appreciated functionality that *connected* their pitch to relevant opportunities and resources, against idea and text generation. The pattern of delegating some execution control while engaging in decision making held across the different information gathering activities described below. Some participants, such as P8 and P19, valued information gathering tasks—P8 describing the process of finding target outlets as an enjoyable "puzzle" and P19 appreciating interviewee identification—highlighting how efficiency-oriented AI features could sometimes conflict with personal fulfillment.

5.1.1 Surfacing opportunities to enable pursuit of personal interests. Most participants saw it as busywork to find and vet different pitching opportunities across news websites, social media, newsletters, public spreadsheets, and so on: "there are like a million places to pitch [...] you got to know if it's accepting at the moment, it changes every day" [P3]. Several participants contrasted this difficulty with the ease of actually writing the pitch, with P12 explaining "It's not the writing that's hard," but rather "trying to figure out what [editors] want."

Delegating this work to AI, as with Pitch Suggest, could allow journalists to identify opportunities which matched their reporting interests with relevant calls for pitches and useful metadata like compensation, outlet circulation, and pitching guidelines (Figure 5). This was arguably the most popular feature across concepts: by connecting journalists' interests to existing editorial demand, it could support efficiency while preserving autonomy over creative decisions. Some participants saw potential for faster transition to the more engaging stage of collaborative work with an editor. P20 captured the ideal role for such systems:
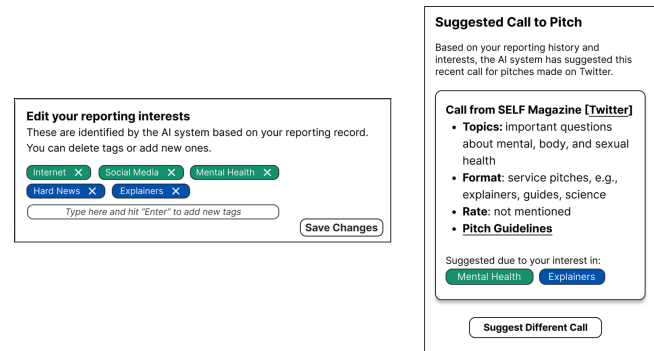

Figure 5: Pitch Suggest would enable participants to share their own writing, pitches, and bookmarks, which the system could use to infer editable tags of their interests (left). These could then be used to match relevant calls for pitches (right).

"If [Pitch Suggest] says 'Hey, SELF Magazine is looking for stories about this [...] it might be something that's relevant to this TikTok mental health study you were looking at recently.' That's the equivalent of introducing two people at a party. Now, it's up to those two people to decide whether they like each other."

Several participants requested filtering controls (e.g., based on word rates) and automatic notifications for opportunities to enhance this connective function. Experienced journalists (e.g., P2, P17) also envisioned AI support for resurfacing their own prior ideas in connection with new pitching opportunities, e.g., from saved links or notes, to enable pursuit of their specific interests. Sharing personal archives to drive these recommendations of opportunities and past ideas, however, also raised plagiarism and impersonation concerns: "How does [the tool] know it's you and not like someone goes in there, and they're like, 'I'm [name]. These are my clips.'? [...] And like trying to use their brain" [P6]. In an ecosystem where a distinctive voice is professional currency, AI that enables such mimicry has material consequences. P7 and P13 expressed a preference for manually specifying their interests rather than having a system infer them from their writing samples, maintaining their decision-making control over the input parameters in the interest of privacy.

5.1.2 Resource recommendations to save time and establish originality. Participants widely accepted delegating the execution of resource discovery to AI systems, as with Pitch Refine's functionality for suggesting sources and datasets (Figure 6a), while maintaining control over which resources to pursue and how to use them. Prior news coverage lookup based on existing pitches could address a significant pain point: "[It] can be kind of tricky knowing if someone else has already published what you want to pitch. You do your best to search with Google and put in the right keywords and figure out if it's already out there, but it's easy to miss stuff" [P8]. Others also envisioned AI-generated suggestions to demonstrate why their story would be timely and original, such as by surfacing relevant trends sourced from Google Trends data,

Current reporting plan

I propose an 800-word news story for MIT Technology Review's website, focusing on the study's implications for how we teach people to check and combat online misinformation. I've conducted preliminary interviews with the study's lead author and will interview a rep from a major search engine company for their perspective on the challenges identified. I can deliver this piece within one week of commissioning.

[Suggest Interviewees]   [Suggest Data Sources]

- **Computational social scientist**: They could offer comments on the study's findings and broader implications
- **SEO expert**: They could explain how current SEO practices might contribute to the prominence of low-quality sources in search.

(a) Participants could hypothetically request interviewee and data source suggestions using Pitch Refine.

Suggest a potential headline I could title my pitch with based on my hook?

Here's a possible headline for your story: "Whispers from the Woods: Citizen Sleuths Unveil the Enigma of Washington's Vanishing Maples"

Suggest a new headline based on my hook in the text.

Here's a potential headline: "Citizen Reports Trigger Urgent Study of Washington's Disappearing Bigleaf Maples"

(b) Participants could hypothetically request headline suggestions using Pitch Assist.

Figure 6: User-initiated engagement for targeted information gathering versus for open-ended text generation.

or unexplored angles from recent news stories. Features for surfacing interviewees and datasets for the reporting plan of the pitch could serve both time-saving and brainstorming purposes, helping identify different stakeholders who might offer perspectives on a story. Participants also requested semantic search capabilities to surface relevant images and visuals for the pitch. Such recommendations could efficiently strengthen a pitch, improving likelihood of acceptance.

Participant responses also show that meaningful control over AI-assisted sourcing requires the thoughtful exercise of professional judgment. Some sought "particular names of experts" [P11], while others recognized that such specific suggestions could reinforce biases, e.g., where "white men [...] pop up first" [P18]. These journalists preferred suggestions of types of sources rather than specific

Customize Critique
Choose criteria for which you would like to receive sentence-level critique from the AI system.

☑ **Accuracy**: Based on references provided by you in-text
☑ **Bias**: Excessive sensationalism or negativity toward a topic
☑ **Clarity**: Ambiguous phrases that can impede understanding
☐ *Custom criteria* *e.g. "potential to add an angle for American outlets"*

Figure 7: Pitch Critic promised granular feedback on both domain-specific and user-defined custom criteria.

names, preserving the open-ended process of considering diversity and bias in their sourcing decisions. Experienced journalists P15 and P20 also emphasized that their field experience would allow them to quickly evaluate AI suggestions for interviewees and datasets, highlighting how the practice of making these sourcing decisions builds the judgment needed to exercise meaningful control.

5.2 Guiding Improvement and Reflection Through Feedback

Participants welcomed AI functionality that could automatically flag potential issues in their pitches (e.g., accuracy, clarity, bias, editorial alignment) while leaving them in control of which changes to make and how to implement them. As with information gathering, this delegation would support efficiency in meeting editorial standards while preserving journalists' autonomy, process, and style. Some likened the feedback to "electronic eyes" [P4] or "a mirror" [P15] on their writing, enabling refinement "but not making decisions for me" [P6]. At the same time, participants warned that AI feedback could subtly constrain their practice in the longer term: by embedding algorithmic bias, devaluing unconventional ideas or stylistic choices, and diminishing the development of analytical skills needed to evaluate the feedback itself.

5.2.1 Domain-specific feedback criteria for efficient scaffolding. Checks for accuracy, clarity, and bias were described as essential but difficult to perform on one's own writing. As P2 explained: "It's very easy for me to edit other people's things [...] when you're proofreading your own stuff or trying to write your own pitch, you're not totally objective because it's you." Manual verification was time-intensive, with P19 noting they had to "triple check what I'm writing based on the papers that I'm citing," and less experienced reporters like P16 often relied on rejections or acceptance from editors to triangulate issues.

Consequently, most participants responded positively to Pitch Critic's domain-specific feedback features, which would flag potential issues against preset journalism values (accuracy, clarity, bias) and user-defined custom criteria (Figure 7). They appreciated how preset criteria might provide proactive reminders: "if there's something that I forgot to ask about, [Pitch Assist] might not suggest it itself, versus [Pitch Critic] will flag it" [P7]. Custom criteria also offered a way to target known weaknesses in their writing, e.g., sentence transitions [P14]. Participants described this structured scaffolding as trading some open-endedness (e.g., with Pitch Assist) for efficiency, while preserving choice over which suggestions to adopt.

Participants acknowledged that verifying AI feedback remained their responsibility. But since such verification was already routine, criteria-based, in-text feedback could help them prioritize attention rather than create extra work. Some did note limitations in how bias might be operationalized within an AI system, given it is an

inherently contextual and subjective construct. For instance, P5 questioned why sourcing-related phrasing was flagged in Pitch Critic, while equally problematic sensationalized language went unnoticed, highlighting how diverging definitions of bias could undermine trust. P13 voiced concern about implicit system bias: "what if the AI is biased […] am I just replacing my bias with AI bias?"

5.2.2 Quick and reflective editing with granular, optional feedback. The ability to exercise meaningful decision-making control over the domain-specific feedback, through its optional and granular nature, was crucial to its acceptance. Pitch Critic promised to highlight specific text segments with short suggestions (typically 4–5 words) and accompanying explanations, which could be quickly accepted or dismissed. Participants saw this as supporting autonomy, originality, and voice: "a check on these various aspects […] in a way you can choose to use or not, rather than generating the idea and the language for the majority of the pitch" [P15]. Explanations and suggestions rooted in user-supplied primary sources (e.g., research papers, other reference material)—"very cut and dry, like suggesting concrete ways of improving without interfering with your actual writing style" [P9]—were perceived to enable more efficient vetting of the system feedback. Several participants indicated they would prefer receiving this in-text feedback at the end of drafting to avoid disrupting writing flow.

The value of such feedback lay as much in prompting reflection as in speeding up edits. P15 described how seeing flagged issues made them "really think about what it is you're trying to convey" and weigh trade-offs in how to phrase findings or frame significance. Experienced journalists like P7 and P20 could quickly discern from the mockup which suggestions fit their voice and what didn't, and anticipated they would still value such assistance as "another editorial viewpoint" [P20]. Yet others worried about mismatches: P8 noted that deliberate stylistic choices, like "sensationalizing […] because that's the vibe [one had] decided," might be mistakenly flagged by the AI and slow them down. P17, a novice journalist, similarly experienced tension between their voice and what they called a "robotic voice" when looking at the mockup, and expressed a preference for feedback from peers who "know my voice." These concerns underscored the need for systems that adapt to individual voice and experience, supporting both efficiency for veterans and style development for newcomers.

5.2.3 Tailoring pitches for unfamiliar outlets in the absence of peer feedback. Participants had mixed reactions to Pitch Critic's editorial feedback feature, which might hypothetically analyze published stories and pitching guidelines to assess appropriateness and suggest improvements for a pitch. Unlike the granular, in-text edits, this feedback was more conceptual, offering alternate news angles, sources, or reporting plans.

While participants saw the potential efficiency gains for targeting outlets, they indicated they would treat this feedback as supplementary rather than authoritative. P15 captured this perspective: "I wouldn't treat it as though it was like a real human feedback […] But given that is so hard to get and almost never happens, it's at least a starting point." Some still preferred human input for conceptual feedback, as P12 explained: "I might ask someone I trust to read over a pitch for me or workshop a lede or something […] Freelancing is kind of lonely, so I would rather talk to another person

about it, and someone that might be my friend." Others stressed that coarse outlet-level analysis might miss specific editors' preferences, limiting precision. Still, participants anticipated the feature would provide a useful baseline for unfamiliar outlets where relationships didn't exist.

P14, who already used ChatGPT to request such feedback in their practice, stressed the importance of knowing "which feedback is worth taking […] what criticism you're willing to accept," from a concept like Pitch Critic, to avoid unproductive pivots. P7's comments highlighted that exercising meaningful decision control in this manner still required the analytical skills the system was designed to shortcut:

> "Any journalist or any science reporter should still know how to figure out that information without relying on [Pitch Critic]. It's a good shortcut once you already are able to do this […] if I want to figure out what STAT is looking for, I would read a bunch of STAT articles […] about similar topics to maybe what I'm pitching them. That's a learned skill, you get better at it the more you do it."

5.3 Kick-starts and Constraints from Ideation Support

Participants were selectively open to AI assistance for idea generation: they welcomed support for peripheral creative elements and overcoming blocks, but insisted on maintaining control over core story ideas. They anticipated this would allow efficiency while preserving autonomy, reflection, and distinctive voice. This reflects a boundary between AI that *kick-starts* the exercise of further causal agency versus AI that substitutes for it or constrains it. Economic pressures and career stage occasionally increased acceptance of more substantial AI ideation, and we note these as they arose.

5.3.1 Supporting peripheral, not core idea generation. Participants welcomed AI support for ideating on peripheral elements of the pitch like potential headlines and email titles (Figure 6b), i.e., "things that don't shape the direction a story is going to take" [P4]. This could help overcome creative blocks without compromising autonomy or short-circuiting the thinking process. Several participants envisioned looking through multiple AI-suggested headlines and titles when feeling blocked and then writing their own. This could provide useful momentum, by making journalists "just feel emotionally that something has started […] look, there's already a thing" [P3]. P20 explained how even low-quality AI outputs could help: "if all of [the AI-suggested headlines] are bad […] all of those are the ways not to do it, now I have a good idea of how to do it."

Resistance grew when AI suggestions presented in the design concepts encroached on what participants considered the core ideas of a pitch, which they felt should remain their own to preserve a sense of fulfillment. Definitions of core ideas varied, complicating any universal design approach, e.g., P1 resisted Pitch Suggest's involvement in selecting papers to cover, P7 opposed the use of Pitch Assist to suggest ledes. P15 described an ambiguous boundary of wanting to have their own "nugget of an idea," but expressed significant comfort with features across the design concepts for suggesting titles, papers, or names of possible interviewees that

could "spark a framing" for a pitch. This resistance stemmed from two concerns. First, participants worried that AI tools could enable widespread homogenization by suggesting similar core ideas to multiple journalists. Second, ceding generation of core ideas could reduce their own effort, which undermined both personal fulfillment and professional relationships. As P15 explained, receiving an AI-generated lede with minimal effort (such as from Pitch Assist) risked losing their "pride that it was my idea," and led to presenting "a false version of myself [...] a more capable one, maybe a more clear one, but not a true one." P8 drew a similar distinction between human and AI collaboration: "I think it would feel a little like I'm cheating if I get [pitch ideas] from a robot, more so than if I get them from a human friend."

5.3.2 Early ideation support as potential constraints on exploration. Several participants welcomed AI's ability to execute work they "couldn't do, or maybe wouldn't do" [P13], like scanning thousands of studies to suggest newsworthy ones, long as they maintained decision control over whether and how to actually pitch them. When describing their overall comfort with different forms of AI support shown in the concepts, P6's foraging metaphor captured this relationship: "I'm grabbing this berry. I'm grabbing that branch [an idea suggested by the tool] [...] But then I get to decide what the meal is [...] I ultimately am the chef of that recipe."

However, other participants emphasized that pitching involved maintaining an open exploratory process, letting ideas emerge through browsing papers, conference materials, and other sources. This exploration offered intrinsic rewards: discovering exciting new topics (e.g., P1, P18); making surprising connections (e.g., P17); and identifying unique angles (e.g., P7). AI-generated ideas, even for peripheral elements like headlines, could undermine this open exploration, even when journalists controlled selection. Timing was critical: concepts like Pitch Assist and Pitch Suggest could generate headlines, ledes, news angles, or even entire pitches before journalists had developed multiple competing ideas, which could "keep brainstorming from being as fluid as it really should be" [P5]. In contrast, processes where journalists first explored broadly and then sought AI input on their ideas—such as playing around with the news angles in Pitch Refine, which required existing drafts—were seen to help preserve the space of possibilities. P17 explained:

> "Once you've got that idea [from Pitch Assist], once you've got that angle in your head, you don't see it in any other way [...] There's something about when it comes from a machine, there's a sense of objectivity that that's the correct answer, that's the right headline. Whereas if I was just brainstorming on a whiteboard [...] a lot of different ideas would feel like they had equal potential."

5.4 Displacement of Effort and Relationships with Text Generation

Participants overwhelmingly rejected AI text generation for substantial elements of their pitches, arguing that ceding execution control (whether for complete drafts or granular elements like ledes and reporting plans) could undermine reflective labor (writing as thinking), distinctive voice, and autonomy over how they directed

(a) Pitch Refine, where users could select specific modifications to apply to their existing lede drafts.

(b) Pitch Assist, where users could provide source material and receive generated lede paragraphs with no prior draft required.

Figure 8: Two approaches to AI-assisted brainstorming: building on existing material with targeted modifications versus generating new material from minimal input.

their efforts. They also worried it could damage authentic relationships with editors who rely on pitches to assess a journalist's capabilities. Even scaffolded and open-ended approaches that might preserve some decision-making control still shifted execution from direct writing to communicating with AI systems, transforming the work in ways some participants found less fulfilling. Nonetheless, even those who valued their reflective labor and voice acknowledged circumstances where practical pressures might justify use and editing of AI-generated drafts, such as exhaustion or fear of rejection. In these cases, acceptance was a compromise driven by circumstance rather than alignment with professional values.

5.4.1 Diminished reflection and distinctiveness with automated text generation. Participants nearly universally rejected the use of AI-generated language directly in their pitches, particularly for substantive elements like ledes or study descriptions. Resistance was strongest toward Pitch Suggest's complete draft generation, which automated the entire pipeline from idea to text, but also extended

to Pitch Assist's paragraph-level outputs and Pitch Refine's modifications when treated as final text rather than brainstorming aids.

This resistance reflected two key concerns. First, AI text generation risked bypassing the reflective work participants valued in writing. P17 explained: "the thinking is in the write of the first draft [...] so if something else writes that for me, then I've lost some of the understanding." Most participants saw writing the first draft as essential for clarifying narrative structure, honing skills, and connecting with their interests, even when pitching was difficult. Second, AI generation threatened to override or distort their distinctive voice and style. P6 articulated the deeply personal nature of voice, beyond its competitive benefits with editors, when confronted with the ability to change it with Pitch Refine's slider-based controls: "I want to be able to have my own voice as a writer and not have words suggested to me to use. And I know I'm saying 'words suggested to me to use', and when it came to like the clarity or accuracy [in Pitch Critic], those were also words suggested to use [...] but this feels different." Several participants also worried that widespread AI use could lead to industry-wide homogenization of writing styles.

P12's cooking analogy captured the link between reflective work, voice, and journalists' sense of fulfillment:

> "There's this line of helping me versus doing it for me [...] Writing, tweaking my voice, or bigger creative idea things feels like doing my work for me, but not in a good way [...] it's like if the robot measured out all the ingredients in a recipe for me versus just cooking it for me. I like cooking. Now I don't get to do any of it."

Some participants distinguished between core creative elements and routine pitch aspects, which required less effort and expression of voice. They expressed comfort with AI generating "mundane bits" [P15] of reporting plans like past reporting experiences [P2, P11, P15], or headlines after "the writer would already know what they want to be writing" [P18].

5.4.2 Loss of autonomy through ceding execution control. At a higher level, text generation functionality was perceived to reduce autonomy over how participants engaged in and directed their efforts, shifting execution control from direct writing to managing AI systems. For instance, rather than allowing them to engage in writing, automated first drafts made some participants feel "like a tool" and "like I'm being used" [P5] by removing control over both initial text creation and the decision to pursue specific story ideas. Substantial text generation also threatened to invert preferred workflows from writing-then-editing to editing AI drafts, with a few participants explicitly stating a preference for tools that "probe my writing instead of putting something out there for me to probe" [P4]. Even partial text generation shifted effort from writing to "having to take time to experiment with and communicate with" tools [P9], or putting effort into aligning outcomes with their own writing voice, an engagement not appealing to all participants. P14 also captured how this hindered decision-making control: since editing pitches for their faithfulness to sources was "way harder than reading the study and then writing your own", AI also risked imposing less efficient execution patterns than journalists would choose autonomously.

5.4.3 Threats to authentic professional relationships. Participants emphasized that relying on AI to generate pitch text could constitute a form of dishonesty with editors, who depend on these materials to assess both a the story idea and a journalist's writing ability. As P20 explained, "if I knew somebody used [Pitch Suggest] all the way [...] I'm not really going to know whether the reporter even understands the pitch, much less would be able to follow it through and deliver on the concept." Some participants noted that using AI to manipulate voice or generate substantial text in pitches could even create implicit pressure to use it for the full news story, since editors often evaluate pitch language as reflective of overall writing voice, especially where journalist-editor relationships don't already exist. P17 connected this sense of misrepresentation to the erosion of authenticity in professional relationships: "tone feels quite inherently linked to personality and creativity and just a sense of care [...] if the tone is being changed with a touch of a button [in Pitch Refine], then it's jeopardizing [those elements]." Participants also worried that these pressures could threaten future collaborations if editors detected discrepancies between AI-assisted pitch elements and the journalist's genuine voice or commitment to the story.

5.4.4 Compromises on values under situational pressures. Despite widespread resistance to automated text generation, participants identified professional circumstances that might justify its use. Situational pressures such as fatigue (P9), anticipated rejection (P3), or the formulaic nature of press release–based stories (P16) created openings to generate first drafts with AI for subsequent editing. As P1 explained, AI-generated first drafts might help when "it's like more of a boring story [...] something like this would be a nice shortcut." Career stage also shaped these calculations: early-career freelancers like P9, P10, and P16 still establishing themselves professionally, prioritized speed and were especially drawn to features that might help produce drafts with less effort, e.g., from Pitch Suggest and Pitch Assist. Yet even when acknowledging efficiency benefits, they recognized the trade-offs, including diminished reflective labor and sense of fulfillment from less hands-on involvement in writing. A few participants, like P6 and P19, ultimately distinguished between short, news story pitches, where automation felt acceptable, and feature-length, investigative ones, which demanded greater executive and decision-making control.

6 Discussion

Our findings reveal how LLM-infused writing tools may reconfigure journalists' agency, and ceding control may or may not align with journalists' professional values. We organize the discussion of our findings in three parts. First, we examine *journalists' boundaries of control* across distinct system functionalities and their design implications. Second, we explore how *causal and identity-based agency can be mutually constitutive*, explaining journalists' resistance to ceding control and the long-term risks of reduced practice. Third, we discuss the *challenges for value-sensitive design*, including individual variation, systemic pressures, and approaches to preserving agency over time.

6.1 Journalists' Boundaries of Control With LLM-Infused Writing Tools

Our findings reveal clear boundaries in how journalists negotiate control over AI during pitch writing. Participants were generally comfortable ceding execution control for information gathering or feedback generation, as long as they retained decision-making authority over how to integrate outputs into their work. Features that surfaced pitching opportunities, suggested sources, or flagged domain-specific issues were widely welcomed as efficiency gains: they reduced tedious labor without encroaching on values central to journalists' craft. However, boundaries tightened around idea generation and text production, which journalists tied to autonomy, fulfillment, distinctiveness, and skill development. Here, AI was acceptable mainly as a kick-start, allowing journalists to then engage in independent execution and decision making. These boundaries, however, were not uniform: early-career reporters sometimes prioritized efficiency to establish themselves, while veterans emphasized critical engagement, expressing their voice, and skill development. Similarly, story type and situational pressures also shaped acceptance at times: for some, AI-generated drafts were acceptable for shorter, formulaic stories or when they faced time constraints, financial stress, and exhaustion. In practice, these boundaries form a jagged frontier, shifting across individuals and contexts.

These boundaries suggest several design implications. First, we argue that efficiency gains that preserve craft values may come more so from connecting journalists to resources that support their practice (e.g., through semantic search capabilities, or pitch opportunity aggregators that parse unstructured data from social media and news sites) than from generating outputs for direct use. This approach can help connect journalists' interests and values with relevant sources and opportunities, rather than distorting their work to fit demand, and aligns to recent work calling for tools that inspire and offer feedback, rather than generate content [29, 40]. Second, professionals' decision-making control in such contexts still needs to be preserved without creating additional cognitive burdens or threats to autonomy. To this end, participants in our study valued optional, traceable suggestions with provenance that they could review individually, rather than automatic, in-place changes. Finally, we also found that some constraints on users' ability to steer or configure the AI system itself may meaningfully enhance agency by focusing attention on editorial decisions rather than managing the AI system. For example, most participants responded more positively to Pitch Refine's scaffolded iteration approach and Pitch Critic's delineated feedback criteria, compared with Pitch Assist's ability to do support iteration and feedback through prompting. The latter offered unconstrained interactions with the AI but also required deliberate focus on prompting, shifting participants' focus away from the text in ways that they deemed unsatisfactory.

6.2 Causal and Identity-Based Agency as Mutually Constitutive

We introduced agency through its causal and identity-based dimensions [5]. This framework further considers identity-based agency in terms of people's self-congruence (alignment of technology use with their values) and self-construction (effects of technology use on values over time). Our findings demonstrate the ways in which the dimensions of agency can be mutually constitutive in practice, where exercising causal agency creates the conditions for sustaining and developing identity-based agency over time.

This relationship explains certain patterns of resistance that we observed. For instance, some participants rejected AI recommendations of specific people to interview, preferring suggestions of types of sources instead. This could preserve the open-ended process of considering diversity and bias, which enables both ethical sourcing for a specific story (self-congruence) and the development of judgment needed to uphold this value over time (self-construction). Similarly, resistance to AI-generated first drafts stemmed from both immediate concerns about expressing voice (self-congruence) and the recognition that drafting practice develops one's distinctive voice (self-construction). Recent studies of AI-assisted qualitative research [41] and UX design [54] also identify similar concerns about developing identity over time. Exercising causal agency in the present helps develop the professional self who can uphold those values in the future.

This mutually constitutive relationship also indicates that prolonged AI use may gradually undermine both dimensions of agency. Exercising meaningful control over any technology requires expertise to evaluate its outputs, and this expertise develops through practice. This need for effortful practice applies even to activities where most participants expressed no concerns about ceding execution control in the short term, e.g., for evaluating outlet suggestions from Pitch Suggest and Pitch Critic, resources from Pitch Refine, or writing feedback from Pitch Critic. But reduced practice may limit development of the expertise needed to evaluate AI outputs, as P7 suggested when considering AI-generated editorial feedback. Indeed, recent work on AI copilots confirms that users often prefer learning by doing, with semi-automated systems supporting learnability better than fully automated ones for creative tasks [33]. Others also worried about cascading dependency, echoing findings from recent work [55]. Using AI-generated text in pitches, which might feel acceptable to some as an efficiency trade-off in a time crunch, would likely require them to extend that reliance to full stories for stylistic consistency. In this way, short-term value trade-offs can constrain long-term control, locking journalists into AI use they would prefer to avoid. Turning to AI for feedback instead of peers or mentors may also reduce opportunities for identity construction through community engagement. We observed that journalists valued community feedback because their peers might understand their voice, draw on shared experiences, and offer connection in otherwise solitary work.

Designers can respond to these concerns by creating conditions for practice and engagement, to help sustain agency over time. Journalists' professional identity develops through routine practices and community interactions, and so we argue that design could help maintain friction and engagement within these processes rather than simply maximizing efficiency [53]. Designers may draw on seamful design philosophy [30] through rendering AI limitations visible and encouraging deliberate involvement in activities where values risk being traded away under situational pressures (e.g., autonomy and voice in writing). Recent work also calls for infrastructure design (e.g., online platforms, communities) that support both disciplined practice and spontaneous exploration in creative work [56].

Focusing AI support on scoped, unremarkable use-cases that support efficiency can also help preserve agency [64]. For example, automatically extracting pitch requirements and metadata from unstructured webpages to match journalists' interests proved to be immensely popular in our study. It could support efficiency without diminishing user involvement in formulating story ideas, ledes, or pitch text. Inverse use cases where AI systems do most of the creative work while people simply feed them inputs can risk reducing users to feeling like "a tool" for the system, as P5 described about Pitch Suggest's first drafts. As generative AI models become increasingly capable, the challenge lies in designing for journalists' contextual needs while maintaining the practices and community connections that make agency sustainable. Future work might also consider illustrating such relationships between agency dimensions—and ensuing design implications—in other knowledge work contexts.

6.3 Challenges for Value-Sensitive Design in Practice

Value-sensitive design (VSD) provides a framework to systematically account for human values in technology design [19]. Our work treats VSD as an approach for designing systems that preserve different dimensions of agency. To this end, our findings build on recent work on value-sensitive design of AI for journalism [35, 48] by establishing the importance of craft values, i.e., autonomy, skill development, cognitive effort, distinctiveness, and personal fulfillment that journalists view as core to their identity. While contextual values like journalistic ethics, efficiency, and newsworthiness often dominate discussions of AI adoption in this prior work, our findings reveal that journalists often hold craft values at par with efficiency needs, especially for activities like ideation and text generation.

However, three challenges complicate VSD in practice. First, these is individual variation in which activities feel professionally meaningful. For instance, some journalists viewed paper searches as mechanical tasks suitable for automation, while others found fulfillment in the discovery itself. This heterogeneity suggests that effective design of LLM-infused writing tools for journalists may entail configurability, allowing them to specify which tasks require support, adjust tool involvement at different work stages, or control how and when suggestions appear, as per their own weighing of different values. Without such flexibility, AI systems risk either automating away activities that individual journalists find professionally meaningful, or preserving tedious work that others might wish to delegate.

Second, systemic pressures can override professional values in practice. We found that most acceptance of AI for idea and text generation occurred when participants envisioned constrained circumstances such as exhaustion, financial constraints, rejection anxiety, or commissions for formulaic stories with quick turnarounds. Acceptance of otherwise agency-limiting AI essentially represented a compromise driven by circumstance rather than professional values, reflecting recent concerns that economic precarity in journalism majorly drives generative AI adoption despite normative concerns [46]. This tension also highlights what Bennett et al. [5] describe as the material dimension of agency, which pertains to the ability of people to act independently or threats to that from coercion,

lack of power, lack of means, and so on. It suggests that VSD must account not only for values reflected by artifacts themselves, but for values in the broader systems into which those artifacts are embedded. One limitation of our work is that the focus on freelance science journalists—who face economic precarity and must continually prove their expertise [2]—likely amplifies these concerns compared to staff writers, although freelancers' autonomy in choosing projects may heighten emphasis on craft values in ways that extend to other creative professionals like writers and designers.

Third, the temporal interplays of agency (Section 6.2) complicate immediate design decisions. Measuring the impact of immediate design choices on long-term agency requires methodological shifts beyond short-term usability studies. Longitudinal studies that track how journalists adopt and appropriate AI systems [42] could reveal whether efficiency affordances cost skill development, whether these shift professional norms, and whether value alignment persists as use becomes routine. Evaluating writing and creativity support tools through the lens of their impacts on users' abilities and processes in this manner still remains at the fringes of HCI [52]. Such studies may also surface value dilution over time [22], i.e., gaps between design intent and actual outcomes. While often representing risk, value dilution could also represent beneficial appropriations. For instance, P8 recognized Pitch Refine's news angle reframing feature as a potential instructional tool for learning how different angles might be written out, shifting it from an efficiency aid to learning scaffold. Understanding how generative AI might reconfigure professional agency requires that HCI researchers practicing VSD attend to the temporal dynamics of value shifts and tensions [18], which might resist conventional measurement.

7 Conclusion

Journalists' agency is crucial to their democratic role, yet the increasing adoption of generative AI in journalism creates risks of value erosion by reshaping how journalists make and execute editorial decisions. Through interviews with 20 science journalists exploring four hypothetical LLM-infused writing tools, we identified nine values spanning craft concerns (e.g., autonomy, fulfillment, skill development) and contextual pressures (e.g., efficiency, editorial alignment) that guide journalists' boundaries of control around AI assistance. While journalists were selectively willing to cede execution control for information gathering and feedback generation, they resisted AI involvement in core ideation and text production, which they view as threatening autonomy, fulfillment, and skill development. These boundaries further reveal that causal agency and identity-based agency can be mutually constitutive: exercising control over decision making and execution builds the judgment, skills, and relationships needed to sustain professional identity over time.

Value-sensitive design offers an approach to navigate these complexities. Our findings suggest that preserving journalists' agency requires designing AI systems that connect journalists to resources and feedback rather than generating content for direct use, maintaining friction in creative processes to support cognitive engagement, and sustaining the communities of practice through which professional identity develops. However, artifact-level design cannot address systemic pressures that might compel intentional value

compromises, nor can short-term evaluation capture long-term impacts of AI use on the development of professional identity. Addressing these challenges will entail accounting for contextual factors that shape individuals' exercise of agency (e.g., economic precarity), and pursuing longitudinal evaluation that tracks gradual shifts in practice, even when these are difficult to measure.

While situated in science journalism's pitch writing, these insights extend to other knowledge-work domains where creative exploration, identity, and ethical commitments intersect with efficiency pressures. Ultimately, preserving agency with LLM-infused writing tools is not just about steerability or configurability, but calls for designing systems that keep humans central, reflective, and in control of shaping their decisions and identities.

Acknowledgments

References

[1] Mike Ananny and Matt Pearce. 2025. How We're Using AI.
[2] Josh Anderson and Anthony Dudo. 2023. A View From the Trenches: Interviews With Journalists About Reporting Science News. *Science Communication* 45, 1 (Jan. 2023), 107554702211491. doi:10.1177/10755470221149156
[3] Martin W. Angler. 2017. Pitching. In *Science Journalism*. Routledge, London, UK.
[4] Franziska Badenschier and Holger Wormer. 2012. Issue Selection in Science Journalism: Towards a Special Theory of News Values for Science News? In *The Sciences' Media Connection –Public Communication and Its Repercussions*, Simone Rödder, Martina Franzen, and Peter Weingart (Eds.). Vol. 28. Springer Netherlands, Dordrecht, 59–85. doi:10.1007/978-94-007-2085-5_4
[5] Dan Bennett, Oussama Metatla, Anne Roudaut, and Elisa D. Mekler. 2023. How Does HCI Understand Human Agency and Autonomy?. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–18. doi:10.1145/3544548.3580651
[6] Oloff C. Biermann, Ning F. Ma, and Dongwook Yoon. 2022. From Tool to Companion: Storywriters Want AI Writers to Respect Their Personal Values and Writing Strategies. In *Designing Interactive Systems Conference*. Association for Computing Machinery, New York, NY, USA, 1209–1227. doi:10.1145/3532106.3533506
[7] Virginia Braun and Victoria Clarke. 2006. Using Thematic Analysis in Psychology. *Qualitative Research in Psychology* 3, 2 (Jan. 2006), 77–101. doi:10.1191/1478088706qp063oa
[8] Virginia Braun and Victoria Clarke. 2024. Supporting Best Practice in Reflexive Thematic Analysis Reporting in Palliative Medicine: A Review of Published Research and Introduction to the Reflexive Thematic Analysis Reporting Guidelines (RTARG). *Palliative Medicine* 38, 6 (June 2024), 608–616. doi:10.1177/02692163241234800
[9] Janet X. Chen, Francesco Vitale, and Joanna McGrenere. 2021. What Happens After Death? Using a Design Workbook to Understand User Expectations for Preparing Their Data. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–13. doi:10.1145/3411764.3445359
[10] Nicole S. Cohen. 2016. *Writers' Rights: Freelance Journalism in a Digital Age*. McGill-Queen's University Press, Montreal Kingston London Chicago.
[11] Nicole S. Cohen. 2019. At Work in the Digital Newsroom. *Digital Journalism* 7, 5 (May 2019), 571–591. doi:10.1080/21670811.2017.1419821
[12] Hannes Cools and Nicholas Diakopoulos. 2024. Uses of Generative AI in the Newsroom: Mapping Journalists' Perceptions of Perils and Possibilities. *Journalism Practice* 0, 0 (Aug. 2024), 1–19. doi:10.1080/17512786.2024.2394558
[13] Mark Deuze. 2005. What Is Journalism?: Professional Identity and Ideology of Journalists Reconsidered. *Journalism* 6, 4 (Nov. 2005), 442–464. doi:10.1177/1464884905056815
[14] Mark Deuze. 2019. On Creativity. *Journalism* 20, 1 (Jan. 2019), 130–134. doi:10.1177/1464884918807066
[15] Paramveer S. Dhillon, Somayeh Molaei, Jiaqi Li, Maximilian Golub, Shaochun Zheng, and Lionel Peter Robert. 2024. Shaping Human-AI Collaboration: Varied Scaffolding Levels in Co-writing with Language Models. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu HI USA). Association for Computing Machinery, New York, NY, USA, 1–18. doi:10.1145/3613904.3642134
[16] Nicholas Diakopoulos, Hannes Cools, Charlotte Li, Natali Helberger, Ernest Kung, Aimee Rinehart, and Lisa Gibbs. 2024. *Generative AI in Journalism: The Evolution of Newswork and Ethics in a Generative Information Ecosystem*. Technical Report. Associated Press. doi:10.13140/RG.2.2.31540.05765
[17] Celia Ford. 2022. *Picking a Publication to Pitch*. Retrieved Jan 26, 2026 from https://www.theopennotebook.com/2022/11/15/picking-a-publication-to-pitch/
[18] Batya Friedman and David G. Hendry. 2019. *Theory*. The MIT Press, Cambridge, MA, 18–57. doi:10.7551/mitpress/7585.003.0003
[19] Batya Friedman, Jr. Peter H. Kahn, and Alan Borning. 2006. Value Sensitive Design and Information Systems. In *Human-Computer Interaction in Management Information Systems: Foundations*, P. Zhang and D. Galletta (Eds.). M. E. Sharpe, New York, 348–372.
[20] Janet Fulton and Phillip McIntyre. 2013. Journalists on Journalism: Print Journalists' Discussion of Their Creative Process. *Journalism Practice* 7, 1 (Feb. 2013), 17–32. doi:10.1080/17512786.2012.657901
[21] William Gaver. 2011. Making Spaces: How Design Workbooks Work. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. Association for Computing Machinery, New York, NY, USA, 1551–1560. doi:10.1145/1978942.1979169
[22] Sucheta Ghoshal and Sayamindu Dasgupta. 2023. Design Values in Action: Toward a Theory of Value Dilution. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference* (2023-07-10) *(DIS '23)*. Association for Computing Machinery, New York, NY, USA, 2347–2361. doi:10.1145/3563657.3596122
[23] Alicia Guo, Leijie Wang, Jeffrey Heer, and Amy Zhang. 2024. Preserving Writer Values in AI Writing Assistance Tools. In *Proceedings of the Third Workshop on Intelligent and Interactive Writing Assistants (In2Writing '24)*. Association for Computing Machinery, New York, NY, USA, 58–61. doi:10.1145/3690712.3690727
[24] Tony Harcup and Deirdre O'Neill. 2017. What Is News?: News Values Revisited (Again). *Journalism Studies* 18, 12 (Dec. 2017), 1470–1488. doi:10.1080/1461670X.2016.1150193
[25] Thomas C. Hayden, Michelle Nijhuis, and SciLance (Firm) (Eds.). 2013. *The Science Writers' Handbook: Everything You Need to Know to Pitch, Publish, and Prosper in the Digital Age*. Da Capo Lifelong Books, a member of the Perseus Books Group, Boston, MA.
[26] David Hesmondhalgh and Andy C. Pratt. 2005. Cultural Industries and Cultural Policy. *International Journal of Cultural Policy* 11, 1 (March 2005), 1–13. doi:10.1080/10286630500067598
[27] Kenneth Holstein, Bruce M. McLaren, and Vincent Aleven. 2019. Designing for Complementarity: Teacher and Student Needs for Orchestration Support in AI-Enhanced Classrooms. In *Artificial Intelligence in Education: 20th International Conference, AIED 2019, Chicago, IL, USA, June 25-29, 2019, Proceedings, Part I* (2019-06-25). Springer-Verlag, Berlin, 157–171. doi:10.1007/978-3-030-23204-7_14
[28] Yuxuan Huang, James Shea, Daniel C. Howe, and Jussi Holopainen. 2025. Lyric Poetry in the Face of Posthumanism: An Analysis of Generative AI-Assisted Poetry Writing. In *Proceedings of the 2025 Conference on Creativity and Cognition*. Association for Computing Machinery, New York, NY, USA, 762–773. doi:10.1145/3698061.3726919
[29] Angel Hsing-Chi Hwang, Q. Vera Liao, Su Lin Blodgett, Alexandra Olteanu, and Adam Trischler. 2025. 'It Was 80% Me, 20% AI': Seeking Authenticity in Co-Writing with Large Language Models. *Proc. ACM Hum.-Comput. Interact.* 9, 2 (May 2025), CSCW122:1–CSCW122:41. doi:10.1145/3711020
[30] Sarah Inman and David Ribes. 2019. "Beautiful Seams": Strategic Revelations and Concealments. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–14. doi:10.1145/3290605.3300508
[31] Maurice Jakesch, Advait Bhat, Daniel Buschek, Lior Zalmanson, and Mor Naaman. 2023. Co-Writing with Opinionated Language Models Affects Users' Views. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (2023-04-19) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–15. doi:10.1145/3544548.3581196
[32] Nikhita Joshi and Daniel Vogel. 2025. Writing with AI Lowers Psychological Ownership, but Longer Prompts Can Help. In *Proceedings of the 7th ACM Conference on Conversational User Interfaces (CUI '25)*. Association for Computing Machinery, New York, NY, USA, 1–17. doi:10.1145/3719160.3736608
[33] Anjali Khurana, Xiaotian Su, April Yi Wang, and Parmit K Chilana. 2025. Do It For Me vs. Do It With Me: Investigating User Perceptions of Different Paradigms of Automation in Copilots for Feature-Rich Software. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, 1–18. doi:10.1145/3706598.3713431
[34] Jeongyeon Kim, Sangho Suh, Lydia B Chilton, and Haijun Xia. 2023. Metaphorian: Leveraging Large Language Models to Support Extended Metaphor Creation for Science Writing. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference*. Association for Computing Machinery, New York, NY, USA, 115–135. doi:10.1145/3563657.3595996

[35] Tomoko Komatsu, Marisela Gutierrez Lopez, Stephann Makri, Colin Porlezza, Glenda Cooper, Andrew MacFarlane, and Sondess Missaoui. 2020. AI Should Embody Our Values: Investigating Journalistic Values to Inform AI Technology Design. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society*. Association for Computing Machinery, New York, NY, USA, 1–13. doi:10.1145/3419249.3420105

[36] Bill Kovach and Tom Rosenstiel. 2021. *The Elements of Journalism: What Newspeople Should Know and the Public Should Expect* (revised and updated 4th edition ed.). Crown, New York.

[37] Philippe Laban, Jesse Vig, Marti Hearst, Caiming Xiong, and Chien-Sheng Wu. 2024. Beyond the Chat: Executable and Verifiable Text-Editing with LLMs. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology* (Pittsburgh PA USA, 2024-10-13). Association for Computing Machinery, New York, NY, USA, 1–23. doi:10.1145/3654777.3676419

[38] Hao-Ping (Hank) Lee, Advait Sarkar, Lev Tankelevitch, Ian Drosos, Sean Rintel, Richard Banks, and Nicholas Wilson. 2025. The Impact of Generative AI on Critical Thinking: Self-Reported Reductions in Cognitive Effort and Confidence Effects From a Survey of Knowledge Workers. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–22. doi:10.1145/3706598.3713778

[39] Mina Lee, Katy Ilonka Gero, John Joon Young Chung, Simon Buckingham Shum, Vipul Raheja, Hua Shen, Subhashini Venugopalan, Thiemo Wambsganss, David Zhou, Emad A. Alghamdi, Tal August, Avinash Bhat, Madiha Zahrah Choksi, Senjuti Dutta, Jin L.C. Guo, Md Naimul Hoque, Yewon Kim, Simon Knight, Seyed Parsa Neshaei, Antonette Shibani, Disha Shrivastava, Lila Shroff, Agnia Sergeyuk, Jessi Stark, Sarah Sterman, Sitong Wang, Antoine Bosselut, Daniel Buschek, Joseph Chee Chang, Sherol Chen, Max Kreminski, Joonsuk Park, Roy Pea, Eugenia Ha Rim Rho, Zejiang Shen, and Pao Siangliulue. 2024. A Design Space for Intelligent and Interactive Writing Assistants. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24)*. Association for Computing Machinery, New York, NY, USA, 1–35. doi:10.1145/3613904.3642697

[40] Zhuoyan Li, Chen Liang, Jing Peng, and Ming Yin. 2024. The Value, Benefits, and Concerns of Generative AI-Powered Assistance in Writing. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24)*. Association for Computing Machinery, New York, NY, USA, 1–25. doi:10.1145/3613904.3642625

[41] Zhe Liu, Jiamin Dai, Cristina Conati, and Joanna McGrenere. 2025. Envisioning AI Support during Semi-Structured Interviews Across the Expertise Spectrum. *Proc. ACM Hum.-Comput. Interact.* 9, 2 (2025), CSCW011:1–CSCW011:29. doi:10.1145/3710909

[42] Tao Long, Sitong Wang, Émilie Fabre, Tony Wang, Anup Sathya, Jason Wu, Savvas Dimitrios Petridis, Ding Li, Tuhin Chakrabarty, Yue Jiang, Jingyi Li, Tiffany Tseng, Ken Nakagaki, Qian Yang, Nikolas Martelaro, Jeffrey V Nickerson, and Lydia B Chilton. 2025. Facilitating Longitudinal Interaction Studies of AI Systems. In *Adjunct Proceedings of the 38th Annual ACM Symposium on User Interface Software and Technology* (2025-09-28). Association for Computing Machinery, New York, NY, USA, 1–5. doi:10.1145/3746058.3758469

[43] Nando Malmelin and Lotta Nivari-Lindström. 2017. Rethinking Creativity in Journalism: Implicit Theories of Creativity in the Finnish Magazine Industry. *Journalism* 18, 3 (March 2017), 334–349. doi:10.1177/1464884915620272

[44] Birgit Røe Mathisen. 2019. Entrepreneurs and Idealists — Freelance Journalists at the Intersection of Autonomy and Constraints. *Journalism Practice* 13, 8 (Sept. 2019), 1003–1007. doi:10.1080/17512786.2019.1640633

[45] Janice Morse. 2004. Purposive Sampling. In *The SAGE Encyclopedia of Social Science Research Methods*, Michael S. Lewis-Beck, Alan Bryman, and Tim Futing Liao (Eds.). Sage Publications, Inc., Thousand Oaks, CA, 885–885.

[46] Allen Munoriyarwa and Mathias-Felipe de-Lima-Santos. 2025. Generative AI and the Future of News: Examining AI's Agency, Power, and Authority. *Journalism Practice* 0, 0 (2025), 1–12. doi:10.1080/17512786.2025.2545448

[47] Sheshera Mysore, Mahmood Jasim, Andrew Mccallum, and Hamed Zamani. 2023. Editable User Profiles for Controllable Text Recommendations. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval* (2023-07-18) *(SIGIR '23)*. Association for Computing Machinery, New York, NY, USA, 993–1003. doi:10.1145/3539618.3591677

[48] Sachita Nishal and Nicholas Diakopoulos. 2025. Values as Problems, Principles, and Tensions in Sociotechnical System Design for Journalism. In *Proceedings of the 2025 ACM Designing Interactive Systems Conference (DIS '25)*. Association for Computing Machinery, New York, NY, USA, 2975–2991. doi:10.1145/3715336.3735717

[49] Sachita Nishal, Jasmine Sinchai, and Nicholas Diakopoulos. 2024. Understanding Practices around Computational News Discovery Tools in the Domain of Science Journalism. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW1 (April 2024), 142:1–142:36. doi:10.1145/3637419

[50] Srishti Palani and Gonzalo Ramos. 2024. Evolving Roles and Workflows of Creative Practitioners in the Age of Generative AI. In *Creativity and Cognition*. Association for Computing Machinery, Chicago IL USA, 170–184. doi:10.1145/3635636.3656190

[51] Zvi Reich. 2010. Constrained Authors: Bylines and Authorship in News Reporting. *Journalism* 11, 6 (Dec. 2010), 707–725. doi:10.1177/1464884910379708

[52] Samuel Rhys Cox, Helena Bøjer Djernæs, and Niels van Berkel. 2025. Beyond Productivity: Rethinking the Impact of Creativity Support Tools. In *Proceedings of the 2025 Conference on Creativity and Cognition (C&C '25)*. Association for Computing Machinery, New York, NY, USA, 735–749. doi:10.1145/3698061.3726924

[53] Advait Sarkar. 2024. AI Should Challenge, Not Obey. *Commun. ACM* 67, 10 (Sept. 2024), 18–21. doi:10.1145/3649404

[54] Prakash Shukla, Phuong Bui, Sean S Levy, Max Kowalski, Ali Baigelenov, and Paul Parsons. 2025. De-Skilling, Cognitive Offloading, and Misplaced Responsibilities: Potential Ironies of AI-Assisted Design. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* (2025-04-25) *(CHI EA '25)*. Association for Computing Machinery, New York, NY, USA, 1–7. doi:10.1145/3706599.3719931

[55] Felix Simon. 2024. *Artificial Intelligence in the News: How AI Retools, Rationalizes, and Reshapes Journalism and the Public Arena*. Technical Report. Tow Center for Digital Journalism, Columbia University.

[56] Ellen Simpson and Bryan Semaan. 2025. Infrastructures for Inspiration: The Routine of Creative Identity Through Inspiration on the Creative Internet. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (2025-04-25) *(CHI '25)*. Association for Computing Machinery, New York, NY, USA, 1–16. doi:10.1145/3706598.3713105

[57] Hari Subramonyam, Roy Pea, Christopher Pondoc, Maneesh Agrawala, and Colleen Seifert. 2024. Bridging the Gulf of Envisioning: Cognitive Challenges in Prompt Based Interactions with LLMs. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24)*. Association for Computing Machinery, New York, NY, USA, 1–19. doi:10.1145/3613904.3642754

[58] Abdullahi Tsanni. 2021. *Making First Contact with Editors*. Retrieved Jan 26, 2026 from https://www.theopennotebook.com/2021/09/21/making-first-contact-with-editors-from-your-desk/

[59] Arjen van Dalen. 2024. Revisiting the Algorithms Behind the Headlines. How Journalists Respond to Professional Competition of Generative AI. *Journalism Practice* 0, 0 (2024), 1–18. doi:10.1080/17512786.2024.2389209

[60] Samangi Wadinambiarachchi, Ryan M. Kelly, Saumya Pareek, Qiushi Zhou, and Eduardo Velloso. 2024. The Effects of Generative AI on Design Fixation and Divergent Thinking. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24)*. Association for Computing Machinery, New York, NY, USA, 1–18. doi:10.1145/3613904.3642919

[61] Sitong Wang, Samia Menon, Tao Long, Keren Henderson, Dingzeyu Li, Kevin Crowston, Mark Hansen, Jeffrey V Nickerson, and Lydia B Chilton. 2024. ReelFramer: Human-AI Co-Creation for News-to-Video Translation. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24)*. Association for Computing Machinery, New York, NY, USA, 1–20. doi:10.1145/3613904.3642868

[62] Justin D. Weisz, Jessica He, Michael Muller, Gabriela Hoefer, Rachel Miles, and Werner Geyer. 2024. Design Principles for Generative AI Applications. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu HI USA, 2024-05-11). Association of Computing Machinery, New York, NY, USA, 1–22. doi:10.1145/3613904.3642466

[63] Allison Woodruff, Renee Shelby, Patrick Gage Kelley, Steven Rousso-Schindler, Jamila Smith-Loud, and Lauren Wilcox. 2024. How Knowledge Workers Think Generative AI Will (Not) Transform Their Industries. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24)*. Association for Computing Machinery, New York, NY, USA, 1–26. doi:10.1145/3613904.3642700

[64] Qian Yang, Aaron Steinfeld, and John Zimmerman. 2019. Unremarkable AI: Fitting Intelligent Decision Support into Critical, Clinical Decision-Making Processes. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–11. doi:10.1145/3290605.3300468

[65] Catherine Yeh, Gonzalo Ramos, Rachel Ng, Andy Huntington, and Richard Banks. 2024. GhostWriter: Augmenting Collaborative Human-AI Writing Experiences Through Personalization and Agency. arXiv:2402.08855 [cs] doi:10.48550/arXiv.2402.08855

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

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

[68] John Zimmerman and Jodi Forlizzi. 2017. Speed Dating: Providing a Menu of Possible Futures. *She Ji: The Journal of Design, Economics, and Innovation* 3, 1 (March 2017), 30–50. doi:10.1016/j.sheji.2017.08.003

A Participant Metadata

Table 2: Years of experience, reporting beat, story type, and the type of target news outlets for study participants.

ID	Experience	Beat	Story Type	Publication Type
P1	0–5 yrs, freelancing	Biology, neuroscience	News, study stories, trend stories	Science-focused, trade publications
P2	20+ yrs, freelancing and editorial experience	Biology, diseases, healthcare	Features, occasional study stories	General interest, science-focused, trade publications
P3	5–10 yrs, freelancing and staff experience	Biology, animals	News stories, features; videos and books as well	General interest, science-focused publications
P4	0–5 yrs, freelancing	Science, nature, environment	Study stories, news, features	Science-focused publications, podcasts
P5	20+ yrs, freelancing (7 yrs) and staff, editorial experience	Earth science, environmental issues, health	Study stories, news, features	Science-focused publications
P6	15–20 yrs, freelancing	Environmental science	Features, news, study stories	General interest, science-focused publications
P7	10–15 yrs, freelancing and staff experience	Life science, technology	Features, occasional study stories	General interest, science-focused publications
P8	0–5 yrs, freelancing	Earth science, life science, climate change, diseases, women's health	Features, occasional study stories	General interest, science-focused, trade publications
P9	0–5 yrs, freelancing	Environmental science, climate, health, wildlife, agriculture, ecology	Features, occasional study stories	Science-focused publications
P10	10–15 yrs, freelancing	Health science, diseases, space	Features	General interest publications
P11	5–10 yrs, freelancing and staff experience	Mental health, health justice, health equity	Features, occasional news stories	General interest, science-focused publications, local news
P12	15–20 yrs, freelancing and staff experience	Neuroscience, physics, other related topics	News, occasional features	Trade, science-focused publications; occasionally general interest
P13	20+ yrs, freelancing (recent, 4 yrs) and staff experience	Life science, ocean science	News, study stories, profiles, features	Science-focused and general-interest publications; occasionally trade publications
P14	10–15 yrs, freelancing	Science, technology, A.I.	Features, podcasts	General interest, trade publications
P15	5–10 yrs, freelancing	Human evolution, fossils, anthropology	Features, news, study stories	Science-focused publications
P16	5–10 yrs, freelancing	Space exploration, astronomy, cosmology, general physics	Features, news, study stories, profiles	Science-focused publications
P17	0–5 yrs, freelancing	Technology, media, healthcare, A.I.	Features	Science-focused publications
P18	5–10 yrs, freelancing and editorial experience	Astronomy, physics, chemistry, geology, science and culture	Features, news, podcasts	General interest publications
P19	5–10 yrs, freelancing	Environment, health	Features	General interest publications
P20	20+ yrs, freelancing and staff experience	Biomedical science, health science, engineering, technology	Features (more recently)	General interest, science-focused publications, local news

B Design Concept Rankings

Table 3: Participants' rankings of design concepts from 1 (most preferred) to 4 (least preferred). Critic = Pitch Critic (feedback), Suggest = Pitch Suggest (end-to-end), Refine = Pitch Refine (iterative brainstorming and writing), Assist = Pitch Assist (open-ended, conversational support). Participants ranking Suggest first all explicitly mentioned preferring its information-gathering features sans automated drafting.

ID	Rank 1	Rank 2	Rank 3	Rank 4
P1	Critic	Refine	Assist	Suggest
P2	Critic	Refine	Suggest / Assist	Suggest / Assist
P3	Critic	Suggest	Refine	Assist
P4	Critic	Refine / Assist	Refine / Assist	Suggest
P5	Suggest*	Refine	Assist	Critic
P6	Suggest*	Refine	Critic	Assist
P7	Suggest*	Critic	Assist	Refine
P8	Assist	Suggest	Critic / Refine	Critic / Refine
P9	Critic	Suggest	Refine	Assist
P10	Critic	Refine	Suggest	Assist
P11	Refine	Assist	Critic	Suggest
P12	Suggest*	Critic	Refine	Assist
P13	Refine	Critic	Suggest	Assist
P14	Suggest*	Critic	Assist	Refine
P15	Refine	Critic	Suggest	Assist
P16	Critic	Suggest	Assist	Refine
P17	Refine	Critic	Suggest	Assist
P18	Critic	Suggest	Refine	Assist
P19	Critic	Suggest	Refine	Assist
P20	Critic	Refine	Suggest	Assist