Technical Report

SummShaper: Empowering Analysts to Tailor Attributed Summaries

Lane Harrison[1*], Melissa Evans[2], R. Jordan Crouser[3], Razieh Fathi[4], Yue "Ray" Wang[5], Hilson Shrestha[1], and Bijesh Shrestha[1]

[1] Worcester Polytechnic Institute; ltharrison@wpi.edu
[2] Vanderbilt University; melissa.j.evans5@gmail.com
[3] Smith College; jcrouser@smith.edu
[4] Whitman College; razie.fathi@gmail.com
[5] University of North Carolina at Chapel Hill; wangyue@unc.edu
* Correspondence: ltharrison@wpi.edu

Abstract: This project aims to explore analysts' behaviors and goals when interacting with attributed summaries. A summary (and its interface) will be more effective if the system understands the analyst's goal: will it be used as a time saver, a takeaway highlighter, an information scent emitter, a mental map, a memory refresher, or a simple answer to a lookup question? We built interactive prototypes for analysts to read, verify, and steer summaries and collected initial data towards understanding the connection between analysts' behaviors and their goals during simulated tasks.

Keywords: analyst behavior; attributed summaries; interactive prototypes; goal understanding

1. Introduction

Recent advances in artificial intelligence have greatly expanded the capabilities of machine-driven document summarization. For many commonly encountered document types, e.g. unstructured text, it is now possible to generate a variety of summaries that reasonably reflect the main components of the source document. Summarization approaches can be categorized across extractive, abstractive, and hybrid methods, each bringing relative strengths and weaknesses. Research efforts are poised to continue making advances in summarization. Summarization speed, quality, and diversity of capabilities will continue to improve, for example techniques for summarizing multi-modal data such as images, diagrams, videos, or technical data. Anticipating this trend, human-machine interaction efforts must integrate summarization capabilities into user-centered interfaces, tools and workflows.

However, the expanding capabilities and system complexities in the summarization space brings with it difficult technical challenges. Summary quality can vary, with problems such as hallucinations, under-/over-emphasized facets, and catastrophic forgetting. Furthermore, summaries themselves cannot also be treated as monolithic– e.g. having one summary for one source document is not a realistic assumption. Instead, summaries might change in form based on context or task. User-interfaces bring it's own set of challenges. These include cultivating trust in summarization techniques and designing usable ways of invoking AI, all while supporting measurably efficient analysis workflows. Interface tools integrating summarization also suffer from a lack of established evaluation measures, methodologies, and frameworks. Taken together, these challenges may significantly slow the adoption of AI-enabled tools and workflows.

In this paper, we describe how we aim to support AI-driven summarization through iterative interface design and controlled experimentation activities. We first develop an interface for displaying information sources (e.g. news documents) and associated summaries with attribution links (see Figure 1). We then embed this interface in an experiment

This technical report reflects work that was conducted as part of the 2024 Summer Conference on Applied Data Science (SCADS) hosted by the Laboratory for Analytic Sciences (LAS) at North Carolina State University.

framework, obtaining responses from $n = 5$ analysts. Feasibility study results and subsequent design sessions revealed an opportunity to move beyond summary navigation to **summary shaping**, where interactive components might empower analysts to tailor summaries to their needs on-demand. We describe a resulting summary shaping prototype, SummShaper, and the opportunities it brings for future studies and systems that support analysts in their challenging and valuable work.

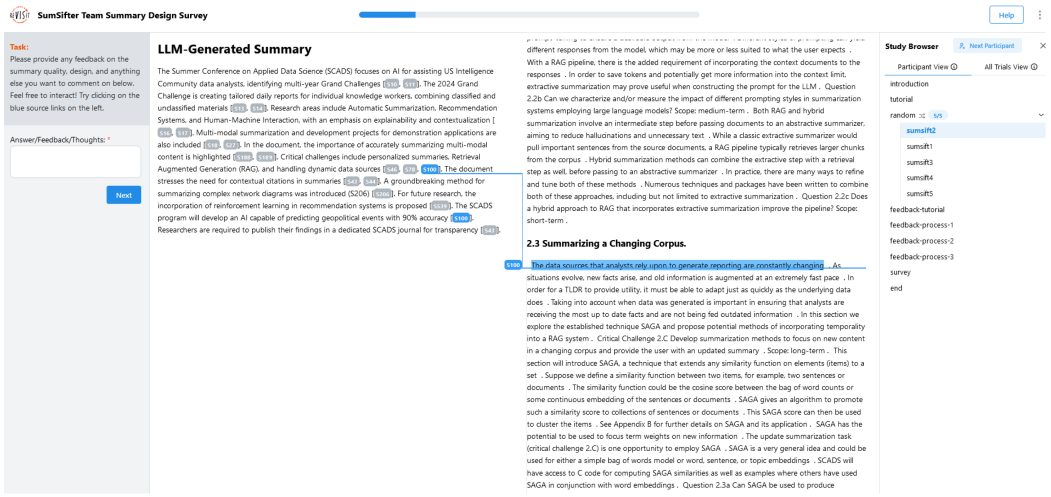

Figure 1. Our initial study examined interactive source document (right) and summary linking (left) via context preserving visual links. Participants received 5 different summary and source pairs and provided feedback in an online feasibility study.

2. Background

The experiments and interfaces in SummShaper draw from prior research spanning summarization, information retrieval, human-computer interaction and data visualization. We briefly review key works that have informed and shaped our approach.

2.1. Document Summarization

Document summarization is a critical area in natural language processing (NLP) that focuses on the generation of concise and coherent summaries from larger texts. Approaches to this problem largely fall into one of two broadly-defined groups:

- **Extractive summarization** involves selecting important sentences, phrases, or sections directly from the source text and concatenating them to create a summary. Early methods relied heavily on statistical techniques such as term frequency-inverse document frequency (TF-IDF) [1] and graph-based algorithms like TextRank [2]. Machine learning approaches further advanced extractive methods by utilizing supervised learning techniques to rank sentences based on their significance [3].
- **Abstractive summarization**, on the other hand, generates new sentences that capture the essence of the source text, often requiring a deeper understanding of the content. This approach is more challenging but aligns more closely with human summarization. Advances in deep learning, particularly the advent of sequence-to-sequence models and attention mechanisms [4,5], have significantly improved the capabilities of abstractive summarization systems. Models such as BERT [6] and GPT-4 [7] have further pushed the boundaries, enabling more coherent and contextually accurate summaries.

Hybrid approaches that combine extractive and abstractive methods have also been explored [8] to leverage the strengths of both strategies. These methods aim to enhance the informativeness and readability of summaries by integrating robust sentence extraction with sophisticated language generation techniques.

Recent trends in document summarization research include the incorporation of reinforcement learning [9], transfer learning [10], and the use of large pre-trained language models [11]. Additionally, there is growing interest in summarizing diverse types of content, such as multimedia, legal documents, and scientific literature, each requiring domain-specific adaptations. The evaluation of summarization systems typically involves both automatic metrics, such as ROUGE [12] and QAGS [13], and human judgment to assess the quality, coherence, and informativeness of the generated summaries. Despite substantial progress, challenges remain, particularly in generating summaries that are factually accurate and contextually appropriate.

2.2. Human-Computer Interaction for Text Data and Search Interfaces

Initiatives in human-computer interaction (HCI) have contributed several models and guidelines for interfaces focusing on text data and search functionalities. This intersection of HCI and information retrieval aims to enhance user experience by optimizing how users interact with text-based information systems. Early work in this domain focused on the development of user-centered design principles, emphasizing the importance of usability and user satisfaction [14]. In 1991, Carol Kuhlthau introduced the Information Search Process (ISP) [15], which underscored the cognitive stages users undergo during search activities. This model has been instrumental in shaping interface designs that support users' evolving information needs.

Building on these foundational principles, more recent initiatives have explored interactive information retrieval (IIR) systems that integrate advanced visualization techniques to aid users in navigating and understanding large text datasets. For instance, Hearst's work on dynamic query interfaces [16] and Shneiderman's development of the Visual Information-Seeking Mantra [17] have guided the creation of interfaces that are not only functional but also intuitive and responsive to user inputs.

The advent of natural language processing (NLP) and machine learning has further propelled this line of inquiry, enabling the design of intelligent search interfaces. These interfaces leverage techniques such as query expansion [18], relevance feedback [19], and personalized search [20] to enhance retrieval performance and user satisfaction. The incorporation of conversational agents and voice-based search interfaces, powered by advancements in NLP, has also emerged as a significant trend, providing more natural and efficient ways for users to interact with text data [21].

2.3. Crowd- and LLM-enabled Search and Summarization

Bernstein et al. explore text-editing interfaces that trigger crowdworkers in the background for tasks such as rewriting sections, grammar checking, etc. [22]. Similar crowdsourcing interface explorations are emerging as models for AI-enabled interfaces, given the potential mapping between intelligent crowdworkers that operate via text prompts and LLMs which can operate similarly [23]. For example, Wu et al. developed AutoGen that organizes LLM agents to perform various tasks, such as code generation, debugging, and question answering [24]. Similarly, Khattab et al. proposed DSPy, a system that can compile a complex question expressed in natural language into a pipeline of LLM prompts, each of which are automatically tuned to improve the final answer [25].

3. Exploring Document Source Attribution Link Interfaces – SumSifter

An early research question that significantly shaped this project was, "How might source attribution links be integrated into summaries to support interactive exploration of source material." Source attribution is an ongoing challenge for LLMs, particularly given challenges such as hallucination which, without links to underlying source material, can be particularly difficult to mitigate. By linking summary claims to underlying sources, users can also investigate sources for deeper exploration of topics of interest.

However, what is not clear is how to actually integrate sources into text. Traditional commercial approaches used summary snippets together with citation-style links [26,27],

which typically open new webpages corresponding to identified sources. Users who wish to validate claims must then search the source document for corresponding information. While this is sufficient for common internet search task scenarios, intelligence analysis often requires deep validation in actual source documents. In other words, any claim in a summary will likely need to be verified in corresponding source documents.

Further complicating the problem is that abstractive summarization, typically done by LLMs, does not by default integrate source information. Sources can also be hallucinated, which may harm trust and necessitate deep analysis, offsetting potential gains of summarization. To explore these issues, we aimed to develop both prompts that would provide source attribution links in an intelligence-adjacent dataset, and interface components that would allow analysts to engage with source attributions and summaries in controlled settings.

An example of the resulting prompt and technique used to generate both summaries and source links is shown in Figure 2.

```
SYSTEM_PROMPT = """
You will be provided with an article in a json format with texts in markdown.
The texts are broken down into blocks and are assigned an id.

You will be prompted to provide a summary of the article.
The sentences in the summary must be attributed to id(s) of the block in the
original article.

You can use markdown within the text block in summary.

Use the following json format to answer.
Do not include sources inside the text block, but instead as a separate list of ids:
{
    "summary": [
        {"text": "Block 1", "sources": ["1", "2"]},
        {"text": "Block 2", "sources": ["3", "4"]},
        {"text": "Block 3", "sources": ["5", "6"]},
        {"text": "Block 4", "sources": ["7", "8"]},
        {"text": "Block 5", "sources": ["9", "10"]}
    ]
}

Do not include any text outside of the JSON format.
"""
```

Figure 2. By using a prompt technique that mimics RAG implementations, we can summarize documents while providing links to source material.

Before the prompt is run, the source documented is separated into chunks with unique source identifiers, similar to RAG approaches [28,29]. The prompt operates on this modified source document as shown in Figure 3, both producing a summary and providing source links in a format that is machine and human-readable. This is the basis for our interface.

```
{
"source": [
    { "id": "1", "text": "Sentence 1: Title of document." },
    { "id": "2", "text": "\n" },
    { "id": "3", "text": "Sentence 2: Overview of the situation." },
    { "id": "4", "text": "Sentence 3: Details of current actions." },
    { "id": "5", "text": "Sentence 4: Official statements or reports." },
    { "id": "6", "text": "Sentence 5: Expert insights or predictions." },
    { "id": "7", "text": "Sentence 6: Analysis of strategies." },
    { "id": "8", "text": "\n" }
]
}
```

Figure 3. The source document is divided into chunks, and numerical identifiers are added to each sentence.

The interface uses the source-links and associated summary to drive the two-pane view in Figure 1. On the left, the generated summary is displayed with source links embedded as citations. The source document appears on the right. An analyst can explore each independently, or click a specific source link to trigger the linked view shown in Figure 1. When sources are triggered in the summary, the source document scrolls to center the associated line, and a context-preserving link is drawn to visually link the two pieces of information [30].

3.1. SumSifter Evaluation Study

With this interface as a baseline, we designed a brief controlled study to bring these components to analysts. Our goals included evaluating the functionality of the interface for source-summary exploration. In some summaries, we also added hallucinated statements that might be noticed and mentioned by the analysts in their exploration.

We developed 5 trials of summaries of varying lengths. Of these, 3 trials included 2 injected hallucinations each. For example, one of the hallucinations was "The SCADS program will develop an AI capable of predicting geopolitical events with 90% accuracy." Another example was "Researchers are required to publish their findings in a dedicated SCADS journal for transparency." The experiment included a brief tutorial and information page. Trials appeared in random order. For each, analysts were invited to comment on their findings. Interaction events and interface state information were logged using information provenance storage techniques. Figure 1 shows an example trial.

3.2. Results

In total, $n = 5$ analysts participated over an approximately 1-week period. Written feedback varied, ranging from 11 to 769 words.

Our initial analysis was of the written comments, which discussed issues ranging from interface suggestions, critiques, reflections on the source data and summarization techniques, and other topics. We use open-coding to categorize several key comments into themes.

Several participants provided feedback on navigation and layout. One participant mentioned that "I really like having the links back to the original source data showing attribution. It does allow me to consider how accurate the summary sentence is with regard to the source. [...] I do think that the links feel a bit distracting to the overall readability of the summary. Maybe something like a hover-over to expand or some other way of hiding the source data when reading the summary would make it overall easier to read."

Another participant noted that "When there are multiple sources assigned to one sentence, it can be challenging to understand which source contributes to the exact portion of the summary, and then it leads to a longer time to read through the actual document." and this participant offered feedback suggesting "What if analysts have the choice to click on the sentence and request a source? I'm thinking maybe that way, visually it's less crowded for analysts to read and you can see what type of sentences triggers analysts to seek out original sources."

Participants also commented on the quality of summaries presented in the trials. One participant noted that "It almost feels like I'm missing a lot of information and, at the same time, not confident that the summary captures the key points accurately." and offered some insights regarding best practices for quality summaries: " [...] I wonder if some way of segmenting the document before summarization (maybe then summarize that section) would improve understanding of the source." Participants recommended maintaining a structure for the summary by suggesting, "For a human reader, keeping in mind these kinds of sentence structure patterns will make the overall summary easier to follow." Observations on data retrieval for summary included comments such as "Assuming a case with some retrieval step used before generating the summary, it would be interesting to think about how to display some insight into which sections were retrieved and why

something was retrieved." These insights helped investigate design spaces regarding summary to source links.

Analysts also identified and mentioned the hallucinations in many instances. As part of the survey, 3 trials included 2 hallucinated sentences each. Observations from participants included comments such as "I noticed several instances where the system generated information not present in the original data, which could mislead users", "summary has some accuracy issues, with citations pointing to irrelevant or redundant excerpts of the text and some summary sentences including information not in the source text. [...]". We take this as confirmation that interfaces providing source links are a suitable avenue for studying issues such as hallucination detection support.

Taken together, these results serve as a preliminary validation of source and summary linking as an approach to support analyst work. At the same time, results suggest several opportunities for improvement and deeper technical work. Later critical feedback sessions with analysts and experts during SCADS revealed two key observations: First, summaries are not necessarily static; they could possibly change based on the context and analyst tasks. Second, summaries could feasibly be shaped in useful ways interactively, given the right UI tools[1].

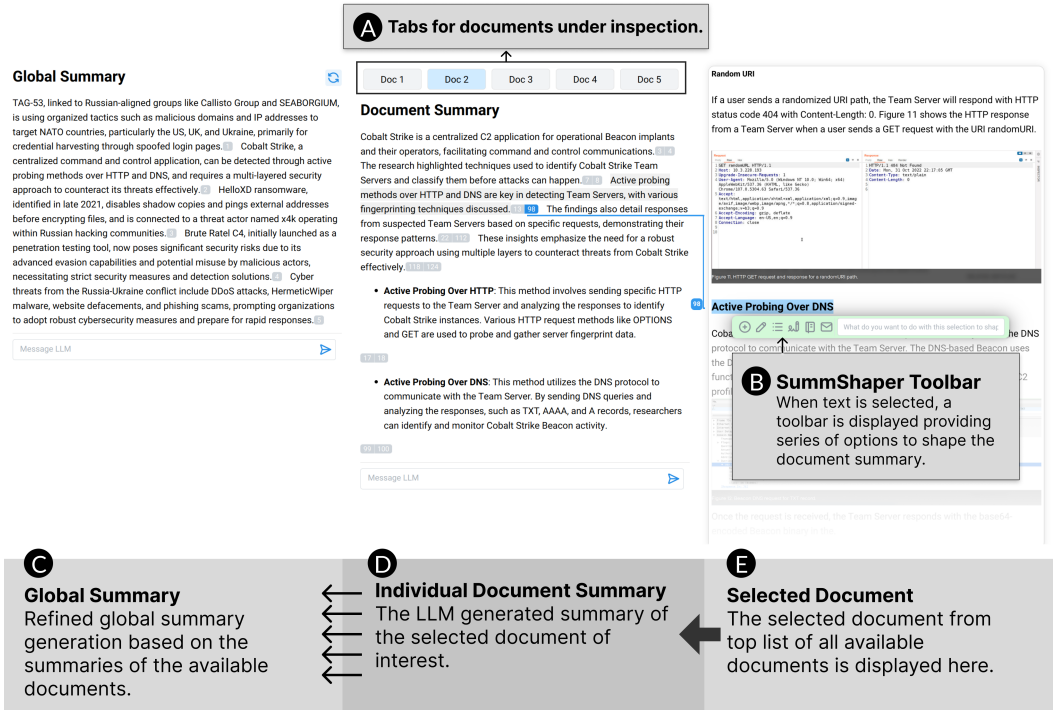

Figure 4. SummShaper: A multi-document summary generation tool that links individual summaries to their source documents. Analysts can explore and tailor both individual and global summaries using the AI-enabled toolbar (B). Sentences in the global summary points to individual documents. Clicking on the document from from the list of available documents (A) displays the document contents on the far right (E). Texts can be selected to shape the Individual Document Summary (D) and eventually the Global Summary (C).

4. From Summary Sifting to Summary Shaping: The SummShaper Prototype

Following the source-attribution link evaluation study and subsequent design sessions, we formulated a second research question to explore: "How might we design interface components and workflows that empower analysts to fluidly shape summaries to meet their needs?". Such an interface, if successful, might be applicable across a wide range

[1] Special thanks to Jonathan Stray for interaction insights in the expert feedback session.

of systems where analysts are (1) engaging with source documents and (2) dealing with LLM-generated summaries of underlying source material.

Operationalizing summary shaping required addressing several fundamental technical and design challenges. For one, the underlying architecture must move beyond the notion of pre-computed summaries, to architectures where summaries can be generated on-demand with custom prompts. Another challenge required rethinking the chat-style interfaces common in most LLM-driven applications, which are laborious to use for shaping summaries. Instead, we sought to identify user-interface techniques that would allow analysts to shape summaries fluidly, with little training, while also providing enough expressive power to remain useful during an active analysis session.

The resulting prototype, SummShaper, is shown in Figure 4. SummShaper builds on the longstanding notion of contextual menus, i.e. menus which appear when users perform certain interactions. However, instead of traditional text-formatting operations, SummShaper replaces these with a custom AI-enabled backend, which is instantiated by familiar actions in a "text-formatting" style toolbar interface. For example, if a user were to highlight a source paragraph of technically-dense text, the SummShaper toolbar Figure 4B would provide a series of buttons, e.g. "Simplify", which would pass the selected text to an LLM with additional selection specific prompts (Figure 5) that would then simplify the text and add it to the summary for further analysis and revision.

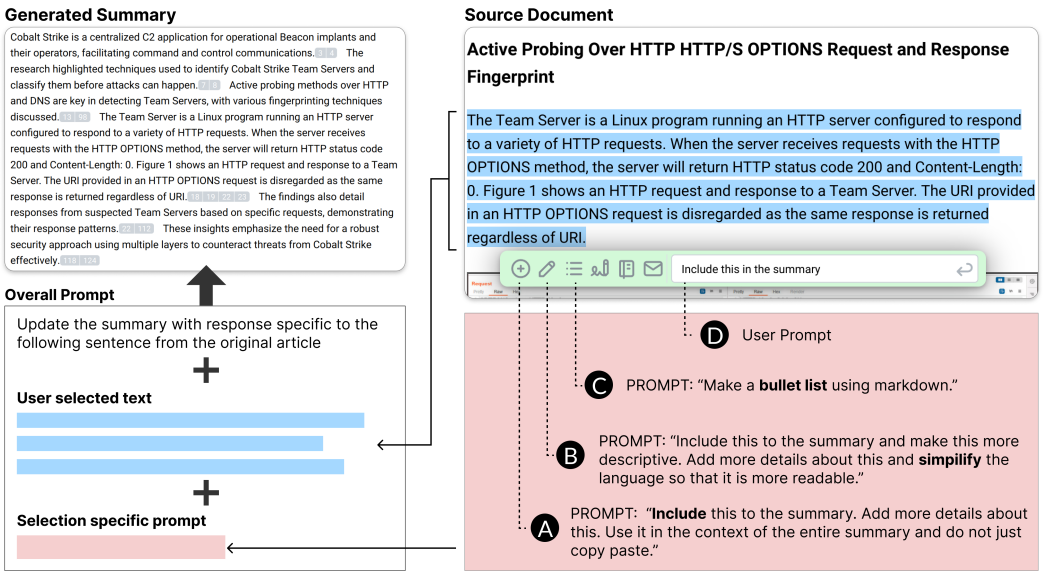

Figure 5. A set of selection-specific prompts is associated with each button in the SummShaper toolbar. These prompts, along with the user-selected text, are combined to create an overall prompt that is sent to the LLM. The response from the LLM is then displayed to the user.

The SummShaper menu is expressive in that any given button could potentially trigger a wide range of high-quality LLM prompts. Text simplification is one example, translation, expanding points, text structure (e.g. bullets Figure 5C), and many more directions are possible. Another distinct advantage to the contextual menu technique is that operations are invoked on specific text in specific windows, which can add valuable information to prompts as they run. For example, the same "simplify" function when a user is in a source document may need to be handled differently when the user is in an LLM-generated summary document. These distinctions can be offloaded to the backend via prompt engineering, creating a more seamless experience for the user.

In addition to the potential for incorporating high-quality knowledge via an easy-to-use interface, SummShaper also provides capabilities for analyst-defined custom prompts, which can be stored and invoked via the menu (Figure 5D). For example, an analyst reading machine-translated source documents from a foreign language could highlight pieces of

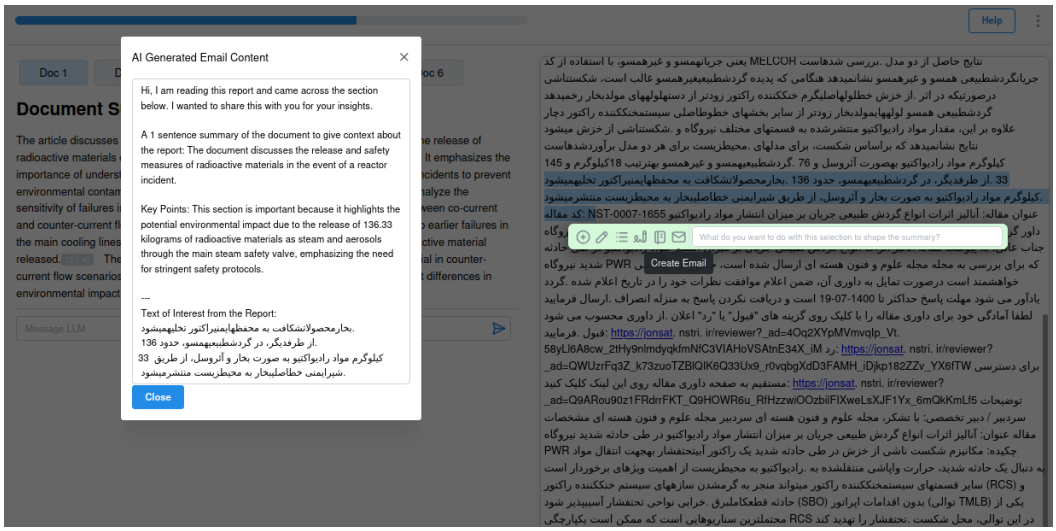

Figure 6. SummShaper also supports cross-language summarization and analyst workflow functionality, e.g. generating an email request for expert translation review.

the text that they wish to have translated by an expert on their team, and use a custom LLM prompt to create a structured request document including source information, links, specific text, and any analyst comments (Figure 6). In other words, SummShaper can potentially streamline some of the laborious information gathering and request work that takes up a disproportionate amount of analyst time.

We present SummShaper as a design prototype which may used in future evaluation studies and as a tool for brainstorming the potential and possibilities of AI with analysts. We discuss opportunities for expanding SummShaper and unique evaluation possibilities, along with implications of the work thus far for other related initiatives.

5. Discussion

Taken together, the results of the studies and design prototypes explored in this work suggest several potential outcomes and integrations with related efforts in summarization and intelligence analysis.

The primary challenge this work highlights is that it remains unclear how to build interfaces that effectively integrate LLMs and other AI-driven capabilities. There are not currently norms, guidelines, or even a wealth of validated systems to draw from. In the early stages of this sub-field, it can be particularly useful for work that maps the difficult work of intelligence analysis to specific interface techniques and workflows. This is one outcome of the work: we explore the intersection of summarization methodologies [31,32], LLMs-as-crowdworkers [23,25], contextual menus [33], and controlled experimentation methodologies for interactive interfaces [34].

Document summarization remains an active and growing area of research itself. This paper sought to draw on those advances to 1) evaluate them with analysts and 2) operationalize them in interface elements that analysts can potentially use as part of their daily work. We envision SummShaper as being an extensible framework for integrating new summarization techniques as they become available. For example, if future research in summarization were to evaluate various prompts for effectively summarizing technical information, SummShaper could potentially integrate such prompts as part of the menu interface. It is also feasible to change out the underlying model, prompts, etc.. to provide new functionality to analysts.

SummShaper also represents a single technique in a design space of many possible interface techniques for integrating and invoking AI. While a contextual menu approach may be effective for invoking AI on specific texts of interest, other interface possibilities could further aid analytical work. Future research might benefit from more granular models

and example datasets that illustrate analytical problems and workflows, which could in turn help spur design activities and a deeper exploration of the design space of AI for the IC.

Finally, evaluation remains a challenge and should be developed alongside the first systems and techniques that integrate AI in analytical contexts. While early work has begun go conceptualize important dimensions of AI in analytical contexts, e.g. trustworthiness and reliability, the difficult work remains of operationalizing these conceptualizations into measurement instruments that can be used to evaluate competing techniques and systems.

6. Conclusions

In this work, we address challenges in integrating AI into user interfaces for analytic work. We develop a user study targeting interactive linking of source and summary documents. From the results of that study, we design and prototype SummShaper, a contextual-menu based approach for invoking AI to tailor summaries on demand. We discuss emerging opportunities and challenges at the intersection of AI, interfaces, and analytic work, and provide recommendations for future studies in the area.

Acknowledgments: The authors would like to thank the SCADS participants and analysts who took the time to provide us with guidance, feedback and ideas. This work was supported in part by NSF#2213757.

References

1. Christian, H.; Agus, M.P.; Suhartono, D. Single document automatic text summarization using term frequency-inverse document frequency (TF-IDF). *ComTech: Computer, Mathematics and Engineering Applications* **2016**, *7*, 285–294.
2. Mihalcea, R.; Tarau, P. Textrank: Bringing order into text. In Proceedings of the Proceedings of the 2004 conference on empirical methods in natural language processing, 2004, pp. 404–411.
3. Kupiec, J.; Pedersen, J.; Chen, F. A trainable document summarizer. In Proceedings of the Proceedings of the 18th annual international ACM SIGIR conference on Research and development in information retrieval, 1995, pp. 68–73.
4. Bahdanau, D.; Cho, K.; Bengio, Y. Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473* **2014**.
5. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. *Advances in neural information processing systems* **2017**, *30*.
6. Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805* **2018**.
7. Achiam, J.; Adler, S.; Agarwal, S.; Ahmad, L.; Akkaya, I.; Aleman, F.L.; Almeida, D.; Altenschmidt, J.; Altman, S.; Anadkat, S.; et al. GPT-4 technical report. *arXiv preprint arXiv:2303.08774* **2023**.
8. Liu, W.; Gao, Y.; Li, J.; Yang, Y. A combined extractive with abstractive model for summarization. *IEEE Access* **2021**, *9*, 43970–43980.
9. Narayan, S.; Cohen, S.B.; Lapata, M. Ranking sentences for extractive summarization with reinforcement learning. *arXiv preprint arXiv:1802.08636* **2018**.
10. Alomari, A.; Idris, N.; Sabri, A.Q.M.; Alsmadi, I. Deep reinforcement and transfer learning for abstractive text summarization: A review. *Computer Speech & Language* **2022**, *71*, 101276.
11. Jin, H.; Zhang, Y.; Meng, D.; Wang, J.; Tan, J. A comprehensive survey on process-oriented automatic text summarization with exploration of llm-based methods. *arXiv preprint arXiv:2403.02901* **2024**.
12. Lin, C.Y. Rouge: A package for automatic evaluation of summaries. In Proceedings of the Text summarization branches out, 2004, pp. 74–81.
13. Wang, A.; Cho, K.; Lewis, M. Asking and answering questions to evaluate the factual consistency of summaries. *arXiv preprint arXiv:2004.04228* **2020**.
14. Shneiderman, B.; Byrd, D.; Croft, W.B. Sorting out searching: A user-interface framework for text searches. *Communications of the ACM* **1998**, *41*, 95–98.
15. Kuhlthau, C.C. Inside the search process: Information seeking from the user's perspective. *Journal of the American society for information science* **1991**, *42*, 361–371.
16. Hearst, M.A. *Search user interfaces*; Cambridge university press, 2009.
17. Shneiderman, B. The eyes have it: A task by data type taxonomy for information visualizations. In *The craft of information visualization*; Elsevier, 2003; pp. 364–371.
18. Carpineto, C.; Romano, G. A survey of automatic query expansion in information retrieval. *Acm Computing Surveys (CSUR)* **2012**, *44*, 1–50.
19. Singh, J.; Sharan, A. A new fuzzy logic-based query expansion model for efficient information retrieval using relevance feedback approach. *Neural Computing and Applications* **2017**, *28*, 2557–2580.

20. Fu, Z.; Ren, K.; Shu, J.; Sun, X.; Huang, F. Enabling personalized search over encrypted outsourced data with efficiency improvement. *IEEE transactions on parallel and distributed systems* **2015**, *27*, 2546–2559.
21. Radlinski, F.; Craswell, N. A theoretical framework for conversational search. In Proceedings of the Proceedings of the 2017 conference on conference human information interaction and retrieval, 2017, pp. 117–126.
22. Bernstein, M.S.; Little, G.; Miller, R.C.; Hartmann, B.; Ackerman, M.S.; Karger, D.R.; Crowell, D.; Panovich, K. Soylent: a word processor with a crowd inside. In Proceedings of the Proceedings of the 23nd annual ACM symposium on User interface software and technology, 2010, pp. 313–322.
23. Wu, Q.; Bansal, G.; Zhang, J.; Wu, Y.; Zhang, S.; Zhu, E.; Li, B.; Jiang, L.; Zhang, X.; Wang, C. Autogen: Enabling next-gen llm applications via multi-agent conversation framework. *arXiv preprint arXiv:2308.08155* **2023**.
24. AutoGen Examples. https://microsoft.github.io/autogen/docs/Examples/. Accessed: 2024-07-19.
25. Khattab, O.; Singhvi, A.; Maheshwari, P.; Zhang, Z.; Santhanam, K.; Vardhamanan, S.; Haq, S.; Sharma, A.; Joshi, T.T.; Moazam, H.; et al. Dspy: Compiling declarative language model calls into self-improving pipelines. *arXiv preprint arXiv:2310.03714* **2023**.
26. Maxwell, D.; Azzopardi, L.; Moshfeghi, Y. A study of snippet length and informativeness: Behaviour, performance and user experience. In Proceedings of the Proceedings of the 40th international ACM SIGIR conference on research and development in information retrieval, 2017, pp. 135–144.
27. Oliveira, B.; Teixeira Lopes, C. From 10 Blue Links Pages to Feature-Full Search Engine Results Pages-Analysis of the Temporal Evolution of SERP Features. In Proceedings of the Proceedings of the 2023 Conference on Human Information Interaction and Retrieval, 2023, pp. 338–345.
28. Gao, Y.; Xiong, Y.; Gao, X.; Jia, K.; Pan, J.; Bi, Y.; Dai, Y.; Sun, J.; Wang, H. Retrieval-augmented generation for large language models: A survey. *arXiv preprint arXiv:2312.10997* **2023**.
29. Pradeep, R.; Thakur, N.; Sharifymoghaddam, S.; Zhang, E.; Nguyen, R.; Campos, D.; Craswell, N.; Lin, J. Ragnarök: A Reusable RAG Framework and Baselines for TREC 2024 Retrieval-Augmented Generation Track. *arXiv:2406.16828* **2024**.
30. Steinberger, M.; Waldner, M.; Streit, M.; Lex, A.; Schmalstieg, D. Context-preserving visual links. *IEEE transactions on visualization and computer graphics* **2011**, *17*, 2249–2258.
31. El-Kassas, W.S.; Salama, C.R.; Rafea, A.A.; Mohamed, H.K. Automatic text summarization: A comprehensive survey. *Expert systems with applications* **2021**, *165*, 113679.
32. Zhang, H.; Yu, P.S.; Zhang, J. A Systematic Survey of Text Summarization: From Statistical Methods to Large Language Models. *arXiv preprint arXiv:2406.11289* **2024**.
33. Benbasat, I.; Todd, P. An experimental investigation of interface design alternatives: icon vs. text and direct manipulation vs. menus. *International Journal of Man-Machine Studies* **1993**, *38*, 369–402.
34. Ding, Y.; Wilburn, J.; Shrestha, H.; Ndlovu, A.; Gadhave, K.; Nobre, C.; Lex, A.; Harrison, L. reVISit: Supporting Scalable Evaluation of Interactive Visualizations. In Proceedings of the 2023 IEEE Visualization and Visual Analytics (VIS). IEEE, 2023, pp. 31–35.

Disclaimer/Publisher's Note: This material is based upon work supported in whole or in part with funding from the Laboratory for Analytic Sciences (LAS). Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the LAS and/or any agency or entity of the United States Government.