Sensemaking Co-pilot

Vision Document

Nivedhitha Dhanasekaran, Jaydev Jangiti, Natasha Joseph, Muyang Xu, Sherry Wu, and Aniket Kittur  
<submission date>  
<version>

# Introduction - Nivi

In the introduction, you summarize, in mostly non-technical terms, the real-world problem your project work addresses, limitations (or absence of) current solutions, how you intend to solve this problem using your idea and envisioned application/method, and the business needs or social motivation behind the project.

1.

# Problem Description - Nivi

Use a heading appropriate for your project context. An example problem description could be: “It takes too long to find an optimal configuration of a Unstructured Information Management Architecture (UIMA) pipeline given a set of possible configurations.” Try to be specific. For example, you could explain the real-world problem statement, why the problem occurs, why the current solution is not a good fit, and ways to overcome these limitations.

This should include a use case when possible! You will be referencing this use case throughout your documents, so think about this.

# Proposed Product/Solution - Natasha

To understand how to utilize LLMs to improve online sensemaking, we will first focus on one domain, learning. Our overall approach will center on an interface that will serve as mode of interaction between the user, browser, and LLM. We’re proposing an end-to-end workflow based on previous work of a browser extension (Fuse) which allows users to categorize and clip information relevant to an overall goal [1]. First, we propose that the system connects to the user’s browser and access all open tabs as the source for the initial synthesis since users usually collect information on a topic from multiple places using their tabs [1]. From there, we plan to extract textual content from the webpages and input this content into the LLM with specific prompts to synthesize important points necessary for what the user is learning (Figure 1). Using the output of the LLM, we’ll structure the information in a concise manner and present it to the user along with references from where the LLM took certain information.

Since the overall goal of the project is to use LLMs to help people learn a new topic, we want to create a tool that can take a breadth of information and summarize it in a manner that can facilitate learning without much manual effort from the user. While Fuse does allow for organization of information to aid with synthesis, it still requires a lot of user intervention to clip relevant images or paragraphs for use later along with annotations to reference later [1]. We hope to eliminate the need for this user input by utilizing the LLM to pull information that will be most helpful for learning. To gather user feedback and improve personalization, we plan to include two interactions (Figure 1) between human and machine by allowing users to highlight information they find useful at the topic or paragraph level [2] on a webpage and like/dislike content synthesized by the LLM to improve the output based on user preferences.

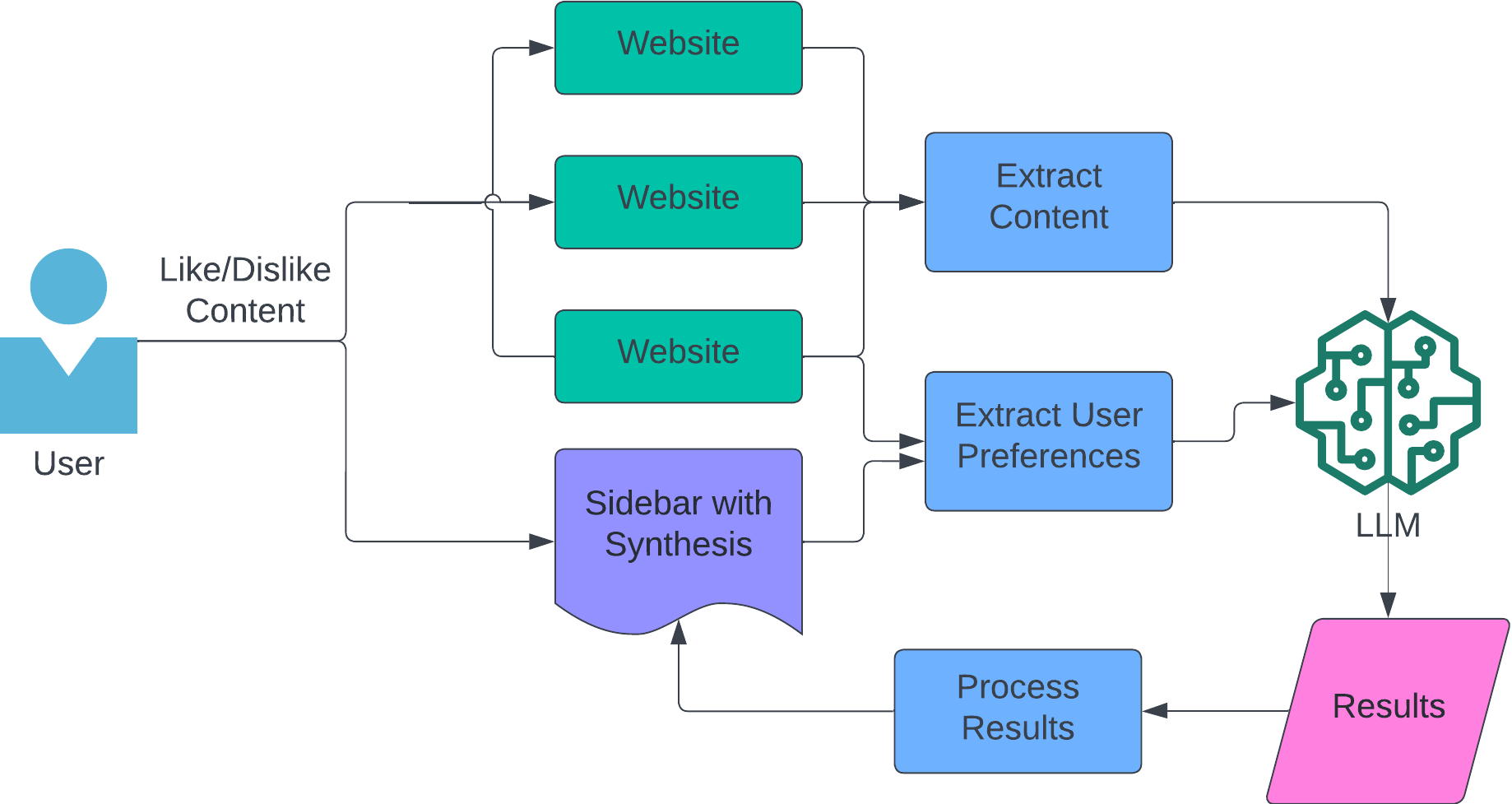


Figure 1: An overview of user workflow with proposed solution

# Scientific Hypotheses - Melissa

1. **Utilizing Large Language Models (LLM) will augment online sensemaking tasks in personal usage scenarios.** This hypothesis posits that leveraging Large Language Models (LLMs) will substantially enhance users' online experiences across various domains, such as online shopping, browsing, event planning, and healthcare consultation. This improvement is anticipated to occur irrespective of the users’ background knowledge or expertise in these specific areas.
   1. **Evaluation metric**: we expect to measure by comparing the relevance of insights generated by users when utilizing LLM versus without LLM support. This comparison aims to quantify how LLMs contribute to users' ability to derive more pertinent and detailed understanding from their online interactions.
2. **Employing a LLM with a carefully crafted prompting strategy can serve as an effective knowledge retrieval tool, enhancing the relevance of the information obtained.** This hypothesis suggests that the design of prompts can impact LLM’s capability to fetch relevant information and synthesize across multiple sources.
   1. **Evaluation metric:** we can quantitatively measure by collecting user ratings of how closely the retrieved information matches their requirement. A higher rating may indicate a better performance of retrieving relevant information. (Reference?)
3. **Integrating LLM into the user's interaction with copilot will improve the operational efficiency of the correction model predictions.** This hypothesis suggests that when users engage with LLMs, they can obtain more accurate information generated by the model to help them better understand the online text information.
   1. **Evaluation metric:** we can measure the time required for users to correct any inaccuracies in the model's predictions. A reduction in correction time indicates an improvement in the model's accuracy and, by extension, an increase in operational efficiency.

Higher-level

1. LLM -> augment sensemaking for personal usage without user’s background information
   1. Metric: ?judge relevance,
2. LLM knowledge retriever → Prompting strategy can help us fetch information
   1. How user satisfies with the model prediction
3. QA baseline???
   1. how long the use needs to fix the model prediction

We can implement online sensemaking to enhance client experience by making it more interactive in different scenarios (i,e. online retail, healthcare, event planning, etc ).

* Evaluation: user-feedback loop???

The proposed solution is designed to effectively retrieve keywords from online literature texts, thereby enhancing the accessibility and understanding of literary content.

The proposed solution aims to enhance sentiment analysis on retrieved literature texts by employing language understanding techniques, and will outperform copilot’s existing sentiment analysis method, demonstrating higher precision and recall in the context of online literature texts. (topic-level, paragraph-level precision)

Prompting strategy -> synthesis

How to interact with synthetic information with LLM (LangChain???)?

The integration of a graph data structure with input text for LLMs is posited to augment the model's comprehension and analysis capabilities by adding layers of derivative meaning to the text. This enrichment can be implemented by several aspects including contextual understanding, relational mapping and inference. The potential of graph data structures can improve the complex reasoning analysis and the precision of semantic analysis. (topic-level, paragraph-level precision)

1loop -> data fit in/ data filtering - undergrade

2 loop -> sensemaking??? Synthesis

Metrics - can we collect more things which are judged as highly relevant like relevance scoring is useful in less time. How much time does it take for user to fix model predictions. How useful is the output judged compared to just LLM baseline.

# Major Features - Jaydev

In this section, you describe in more specific terms how the system features you plan to implement from the proposed solution to the problem (e.g., a list of major features or components of your solution). It is basically a short version of the design document. Consider numbering each major feature, as you can reference back to them in your Design Document. This is really important from a "traceability" perspective as it's easy to miss major features in future documents.

* Intuitive user interface for using the co-pilot
* Detection of contextual/inference signals
* Better online information retrieval using constraints for recommendation/manipulation - Getting structured output from LLM
* User Study - User evaluation

# Scope

Final deliverables for the capstone will include an interface that allows a user to use an LLM to synthesize content based on open browser tabs. Through this interface the user will be able to provide feedback to the model in the synthesized information or in the web page itself. The solution will be deployed in the web browser as an extension for real-time synthesis of information. Our work will focus on how to use the LLM to make sense of the information that is currently available and not on the information foraging aspect of sensemaking. Although foraging could help users identify sources from where to begin learning, including retrieval could prevent us from achieving the primary goal of creating a useful tool within the time constraints.

# Timeline - Nivi

Include a proposed schedule of the proposed work. Take into consideration the timeline of the Capstone sequence.

# Terminology, Definitions, Acronyms, and Abbreviations - Everybody

Include all definitions, acronyms, and abbreviations necessary to understand your solution easily. You can use a table to organize your definitions if necessary. For example, define the NFR’s, the metrics that are going to be used in the project, words associated with Data Structures (e.g., what is a feature vector?), terms used in the project (e.g., what is meant by a classifier?), etc.

Large Laguage Model (LLM):

# References - Everybody

[1] Kuznetsov, A., Chang, J. C., Hahn, N., Rachatasumrit, N., Breneisen, B., Coupland, J., & Kittur, A. (2022, October). Fuse: In-Situ Sensemaking Support in the Browser. In Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology (pp. 1-15).

[2] Pirolli, Peter & Card, Stuart. (2005). The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis.

# Reflection - Everybody

Use this section to write a brief description of what you learned in the process of making this document: what will we do differently next time, what we learned from working in a team, etc. Then, reflect on the decision-making process when making this document. Reflection points to consider:

1. In what ways did creating a vision for your project help reduce ambiguities and establish a shared understanding among all stakeholders in a project?
2. How was the scope and vision of the project formulated? What factors did you consider in this process?
3. Did you consider any alternate approach but had to let go? Why did you pick the current approach over the discarded one?
4. How did you resolve team conflicts while coming up with the vision, if any? Were all possibilities thoroughly examined?
5. How effective was the team collaboration and communication while making this document? What did you learn about teamwork and communication?
6. How did you ensure that everyone was on the same page?
7. Were there any turning points in the process of creating the vision document? If so, please explain.
8. What would you do differently next time?

To gain a deeper understanding of the document's requirements, our team has scheduled two stand-up meetings with our mentor. During these sessions, we will meticulously review each section to clarify the expected contents and ensure we meet all specified criteria. After the meetings, we conducted a swift catch-up to understand each team member's schedule and, using Trello, flexibly assigned document sections to individuals based on our current workloads. Assigning clear roles and responsibilities helps in managing deadlines by implementing the workflow, reducing overlaps and gaps in contributions during the teamwork.

We schedule the regular check-in meetings with mentor to track progress and address any issues properly. Besides, we utilize shared platform such as google drive and slack to track our documentation progress, ensuring easy access to information for all team members.

We did encounter some turning points during the preparation of the vision document. One such turning point arose when we realized our initial draft of hypotheses was overly specific and concrete. This level of specificity was not conducive to the foundational nature of our project, as it limited our scope for exploration and adaptation.

For the following project process, we plan to adjust our weekly schedule in more flexibility to adapt to unexpected changes without derailing the project timeline. In future endeavors, we aim to engage in more frequent discussions with our mentors to ensure our process remains on track.

# Appendix

This section contains any additional information you’d like to preserve in this document for context. For example, consider having a Glossary or any additional materials you discovered or created in the process of making this document.

—-----------------------------------------------

# Writing, Style, and Formatting

### Overall writing requirements

Make sure your writing is brief and easy to understand. 1 to 3 paragraphs per section. The entire document (not including references and appendix) should be about 2 to 5 pages, although this is not a strict requirement and is adaptive to the variety and differences among projects. Please take time to edit and proofread your work before submitting it.

Proper citation practices must be followed, giving credit to original sources of information and ideas. All in-text citations should be clearly marked and correspond to a detailed reference list at the end of the document. This list should include all the sources cited within the text, formatted according to a consistent citation style.

### As always, if you produce subsections

Make sure that you use the proper sub-heading style.[[1]](#footnote-0)

### The same goes for Sub-sub-headings

This is important because the documents you produce may be read by people who are not close collaborators and for whom a well-structured document is helpful to understand things. Also, remember to cite the things you use [C&I, 2016].

1. It may not appear necessary at first but it is part of learning how to communicate your work. Sometimes you may want to add auxiliary information into footnotes such as this one. Examples include technical things like URLs, reference numbers of any kind or citations to papers and external documentation. [↑](#footnote-ref-0)