University of Toronto
Faculty of Applied Science & Engineering
APS112: *Conceptual Design Specification (CDS)*

Team #	030	Tut #	0119	Date	03/18/2024
TA	Jones Law	CI	Patricia Peters	EM	Steven Thorpe
Client Name	Andrea Kassner				
Project Title	Automated Literature Review				
Prepared By (Names & Student #s*)	Sameha Tasnim Nathaniel Ng Hisham Kaleem Nameera Ahmed Vidya Mukherjee Adam Abreha				

Disclaimer:

The design presented in this report has been prepared by first-year engineering students at the University of Toronto and does not present a Professional Engineer design. A Professional Engineer has not reviewed this design for technical accuracy or adequacy. The recommendations of this design may not be implemented in any way unless reviewed by a licensed Professional Engineer where such review and approval are required by professional or legal standards; it is understood that is the responsibility of the recipient of the design to assess whether such a requirement exists. This report may not be reproduced, in whole or in part, without this Disclaimer.

//

Engineering Strategies & Practice

Executive Summary

Dr. Andrea Kassner, a Senior Scientist at the Hospital for Sick Children, requested an autonomous, digital system that can perform systematic literature reviews, create summaries and simple PowerPoint presentations for a given medical research subject. Systematic literature reviews follow a pre-established protocol and use rigorous and transparent methods to synthesize a comprehensive, relevant, and unbiased review.

There is a gap in current literature review systems as they are only able to summarize articles that the user must input manually, while the client requires a system that can search through an entire online database such as PubMed and Google Scholar and analyze and summarize all literature related to the user's research. Furthermore, the system must also be able to take in literature from the user's personal library on Zotero.

There is a need for a system that will reduce time and labour when analyzing and reviewing articles. This will help accelerate research, help users prepare for medical Journal Clubs, aid in teaching, and aid in reviewing grant applications.

The system will only analyze English medical literature, will only be available to those within Dr. Kassner's lab, and must be secure and password protected. The system must be compatible with Windows OS, must be able to cite in the IEEE style and may also give the option to cite using MLA or APA styles. The solution can incorporate pre-existing AI and language models. The system aims to be fast, easy to operate, secure, and should be accessible to the whole laboratory. Furthermore, the design should be able to take in input from the client about citation style, and length of the desired summary.

After generating 89 ideas, the team narrowed down the final design to three different solutions using feasibility checks, multivoting, morph charts, graphical decision charts and pugh charts. The first solution incorporates independently developed technologies along with commercially available technologies such as IBM Watson NLU, GPT-4 and official APIs. The second solution focuses on independently developing and coding the components of the design using Python, different APIs and libraries. The third solution only incorporates pre-existing technologies such as Scrapy, Perplexity, PopAI, Pitch and Duo Security. After careful discussion it was determined that the first solution was the best as the IBM Watson NLU would allow for a broader range of inclusion criteria and a more rigorous filtering of literature to remove bias.

In conclusion, the proposed solution aims to reduce time and labour by autonomously performing systematic literature reviews and producing editable summaries and PowerPoint presentations that can be used in Journal Clubs, for teaching and for reviewing grant applications. In terms of the next steps, the team will meet with the client on April 11th and review the CDS. After revisions, the team will work towards implementing the measures of success by beginning to prototype and test. In terms of long-term goals, the team will be preparing for the final presentation.

1.0 Introduction

A systematic literature review is a methodical process of gathering, evaluating, and summarizing relevant research on a specific topic to provide a comprehensive, relevant and unbiased overview of the current knowledge in the field (Figure 1). Conducting such reviews is time-consuming and effort intensive. Thus, Dr. Andrea Kassner of the Hospital for Sick Children's research team, has tasked the group with automating this process along with producing PowerPoint presentations. The program will make use of generative AI. Refer to Appendix A for a glossary of terms.

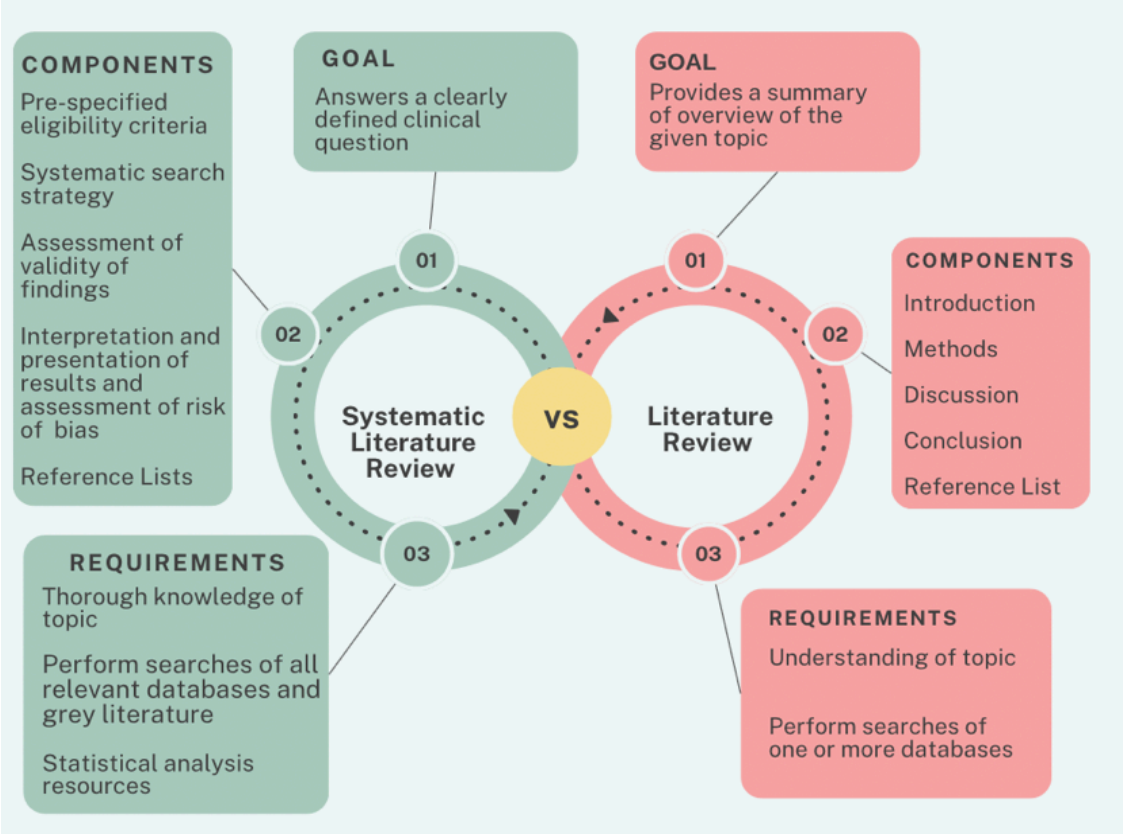


Figure 1. Differences between literature reviews and systematic literature reviews.

2.0 Problem Statement

The client, Dr. Andrea Kassner, requires an automated system that can analyze medical literature from PubMed, Google Scholar and her personal library stored on Zotero. The system must perform a systematic literature review and output a summary of all the analyzed literature, and two PowerPoints of the summarized information. Current technologies such as PopAi [1] are only able to analyze one article at a time and cannot search through an entire database.

2.1 Gap

There is a lack of a virtual, autonomous system that can accept articles from the user's personal library, search through an online database for literature related to the user's research and analyze the literature to produce a summary of all the articles, and PowerPoint presentations.

2.2 Need

There is a need for an autonomous system that can input literature and output summaries, and PowerPoint presentations to help save time and labour when preparing for medical Journal Clubs, teaching, and reviewing grant applications.

2.3 Scope

After consulting the client, the team decided that the system will only accept literature from PubMed, Google Scholar and the user's personal library stored on Zotero. The system will only analyze medical literature in the English language. The system will only be available to Dr. Kassner's lab and will not be open to the public. The team's design can incorporate pre-existing generative AI and language models.

3.0 Service Environment

The service environment of the project is limited to technology as most of the project requires an interaction between the individual and the device. The service environment is described in Table 1.

Table 1: Description of the Physical, Virtual, and Living Environments

Classification	Details
Physical	Compatible devices: ● Laptop ● Computer (PC)
	General hardware specifications of the device used to run the system (minimum specifications): ● 16-32 GB of RAM ● 12th Gen Intel ® Core™ i5-12400F, 2500 Mhz, 6 Core(s), 12 Logical Processor(s) (D. Sare, personal communication, March 19, 2024) ● 3060 RTX GeForce GPU (D. Sare, personal communication, March 19, 2024) ● Storage space on computer [2] ● Microsoft Windows 11 Pro
Virtual	Requirements for the system: ● Wireless capabilities

	○ Wi-Fi (2.4 GHz & 5 GHz) [3] ● Compatible with Windows OS
	Software integrated into the system: ● Microsoft ○ Word ○ PowerPoint ● PubMed to obtain articles ● SickKids personal library ○ Zotero
	Threats to software: ● Viruses ● Spyware ● Fake security software
Living	Operators: ● Researchers of Dr. Kassner
	Malevolent beings: ● Hackers

4.0 Stakeholders

While the design will be used by Dr. Kassner and the four members of her lab [4], the following table lists stakeholders in descending order based on how they are impacted by the design and how they influence it. Refer to Appendix B for a graphic comparing the influence and impact of the stakeholders.

Table 2: How Stakeholders are Impacted and their Influence

Stakeholder	Impact	Influence
Academic Researchers	● Citation count of researchers will increase. ● System will be able to analyze more articles.	If literature is not cited properly, legal issues are caused, making the researchers highly influential.
Natural Sciences and Engineering Research Council of Canada. (NSERC)	● Facilitates reviewing grant applications. ● Will likely increase the number of grant applications.	Researchers requiring grants must concisely submit appropriate summaries [5] and proposals for research, which highly influences the team's future designs.

Journal Clubs	• Design reduces the effort and time required to make research-based PowerPoints.	Main use of PowerPoint is for Journal Clubs, making them highly influential in the development of presentations.
Education Institutions	• Creates PowerPoint presentations at different linguistic levels. • Researchers are more likely to present their research and educate students.	Client noted that teaching will be an added feature rather than a main priority, causing them to have a lower influence.

5.0 Detailed Requirements

Functions, objectives, and constraints for the design are outlined below.

5.1 Functions

The autonomous system should be able to conduct the following main function, and the other functions that enable or stem from it. These were generated using the functional basis and black box method (Appendix C).

5.1.1 Primary Function:

- Transform inputted data in the form of articles from PubMed and Google Scholar into summaries and PowerPoint presentations.

5.1.2 Secondary Functions:

To enable the primary function, the design should:

- Accept general keywords and MeSH (Medical Subject Headings) terms from the user.
- Retrieve data from the PubMed and Google Scholar database using inclusion criteria and search strategies.
- Allow manual removal of data by user to narrow down information.
- Condense and organize relevant information into digestible bullet points to form summaries.
- Convey data through Word documents and PowerPoint presentations.

The following results from the primary function:

- Allow copying/pasting of data into other applications.
- Export citations of referenced literature into separate Word documents.

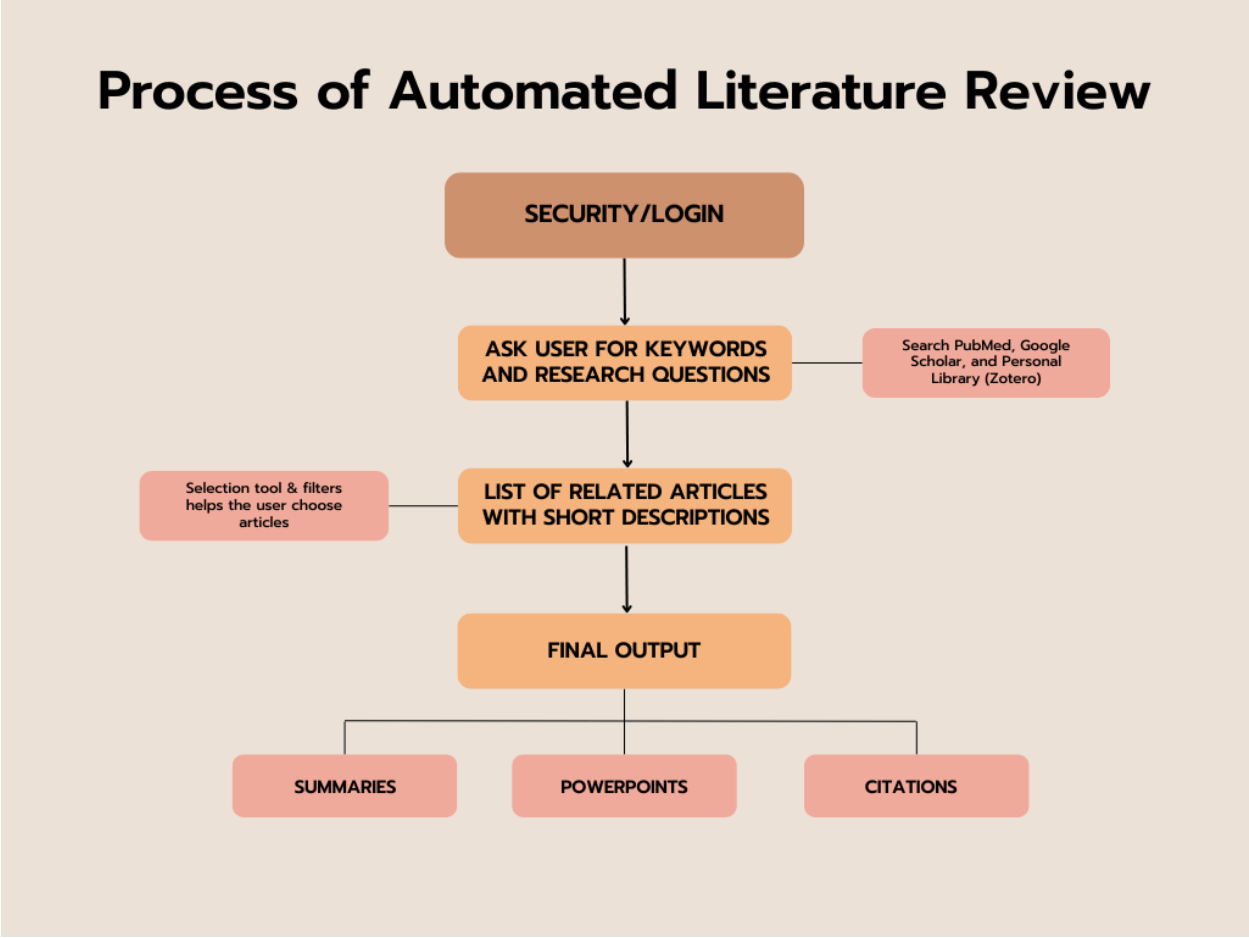


Figure 2. The process of the automated literature review.

5.2 Objectives

The primary identified objectives for the project, along with their associated goals/metrics, are outlined below.

Table 3: Objectives and Associated Goals/Metrics/Justifications

Objective	Goal	Metric	Justification
#1. The solution should be easy to operate	The following ranges should be achieved: • **5-8** clicks for data input • **2-3** clicks for processing • **2-3** clicks for obtaining/downloading output	# of clicks [6]	Client requested the solution to be easy/intuitive to use. The cited paper indicates minimal steps/effort in the form of # of clicks to be a good indicator of the useability of software. Established ranges are based on tests conducted by team

			members on generative AI softwares (QuillBot, PopAI, etc.).
#2. The solution should be versatile in its input/output	The following ranges should be achieved: • **1-3** options for database selection (digital collection from which articles are used) • **8** options for text complexity (level of difficulty in reading the text) • **2** options for PowerPoint format (color scheme/layout of presentation) • **1-3** options for citation, paper, and PowerPoint export format (file type items are exported to)	# of input/output options	The ranges were established by the client.
#3. The solution's output should span a broad range of comprehension levels	The Flesch-Kincaid Reading Ease - a scale indicating the difficulty of reading and comprehending the text - should have the ability to range from 60-0 [7]	Flesch-Kincaid Reading Ease	The client has requested that the summary should range from high school to doctoral-level comprehension (score between 60-0) [7]
#4. The solution should be fast to use	The entire process from loading system to receiving/downloading output should take ≤ **5** minutes	# of minutes	The client requested the solution to be fast to use. The time bound was obtained from the client.
#5. The solution should be secure	• System should have **3** account types with varying authority (Admin, User, Guest) • Password entropy - a measure of the complexity/unpredictab	# of account types, Entropy (in bits)	Account types were obtained from the client. Password entropy goal is established by Anna Szczepanek, PhD: "For non-vital accounts, 25-30 bits of entropy are enough. For more important accounts, aim

	ility of the password - should be **60-80**+ bits		for 60-80+ bits of entropy" [8].

5.3 Constraints

Constraints are classified by source into the following categories: clients and regulations.

Table 4: Design Constraints from the Client and Regulations

Category	Constraints
Client's	Use PubMed and Google Scholar as its database
	Not have a physical component except for the desktop itself
	Must be password protected as only members of Dr. Kassner's lab will be able to use it.
	Two editable PowerPoint templates are created. One with a black background with yellow or white text and one with a white background and black font.
	PowerPoint is formatted in bullet points
	Citation must be in IEEE and potentially APA or MLA format depending on the summary
	Compatible with Windows OS
Regulations	In the case that the system uses generative AI, it must abide by Artificial Intelligence and Data Act (AIDA) which was tabled as part of Bill C-27 (Appendix F).

6.0 Generation, Selection and Description of Alternative Designs

This section introduces the processes and tools used to create and combine solutions to generate three alternative designs. The team generated 89 ideas in total (Appendix G). The alternative designs are justified with their performance against the key objectives (Appendix I).

6.1 Idea Generation Process

The team decided the main groups of ideas should align with the objectives and be input/output versatility, security, ease of use, and speed (Appendix G). A morph chart was created based on these objectives (Appendix H).

The team generated ideas using methods such as blue sky thinking, free brainstorming, lateral thinking, SCAMPER, and magic solutions (Appendix G). Each team member came up with a complete solution made up of the smaller ideas satisfying each objective (Appendix H). Refer to the diagram below to view the team's idea generation process in chronological order.

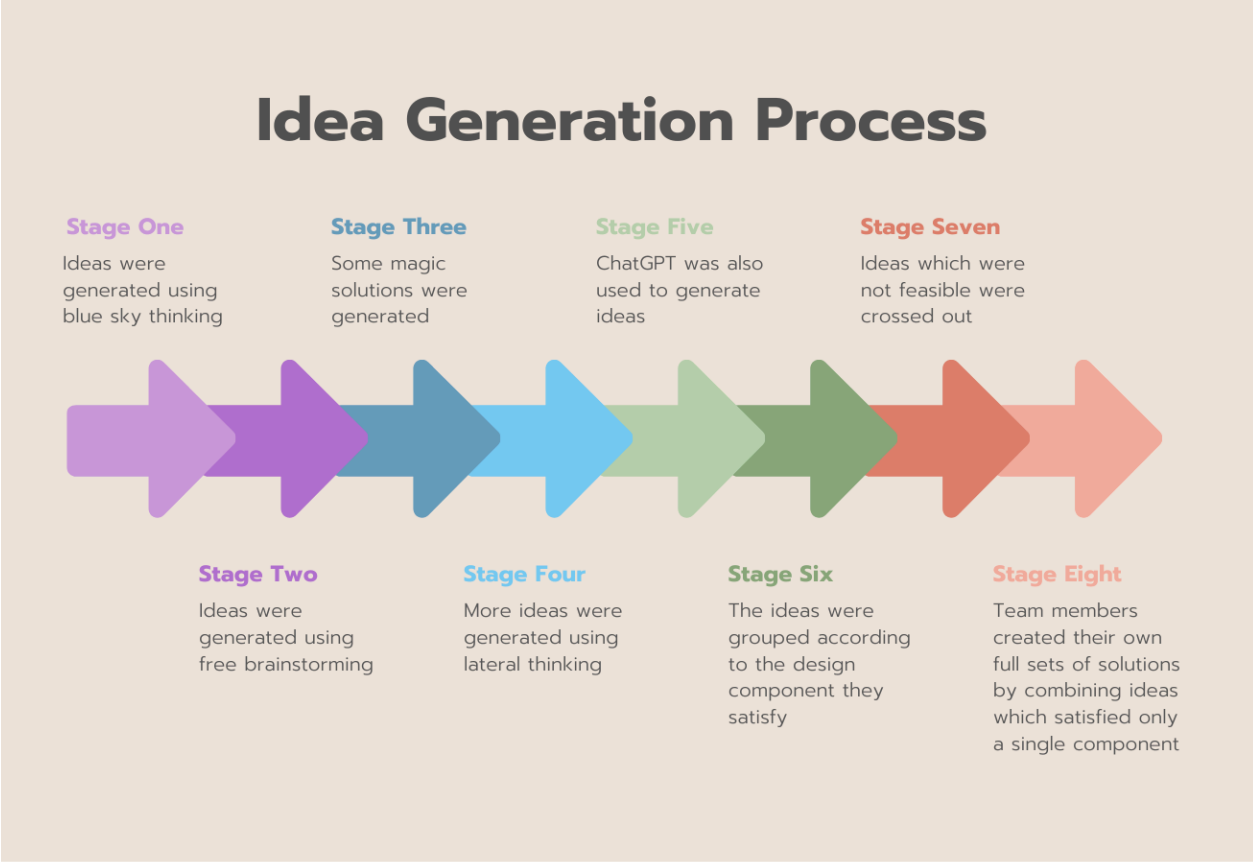


Figure 3. Step-by-step idea generation process.

6.2 Alternative Design Selection Process

The team first eliminated duplicate ideas. Thirteen ideas that were infeasible and did not align with the functions, objectives and constraints (FOCs) were then eliminated (Appendix G). Multi-voting allowed members to vote for seven different ideas (Appendix G). Twenty-four ideas having ≥1 votes were further considered. Each member was responsible for creating one alternative solution which incorporated ideas from the list (Appendix H). Three final alternative designs were selected from the graphical decision chart, which demonstrates how solutions perform against the highest priority objectives: versatility and ease of use (Appendix I).

6.3 Alternative Design Descriptions

This section describes the three proposed designs, their specifications, and how they satisfy the FOCs. Refer to Table 3 for the list of objectives.

6.3.1 Solution 1: Commercial combined with Self-developed Technologies

This solution focuses on combining commercially available and independently developed technologies. The solution follows a streamlined process identical to Figure 2 with the added feature of a literature bias detection framework that will further differentiate collected articles based on their objectivity (key difference from the other solutions). Additionally, the programming of the solution will be outsourced to an experienced software developer that will be able to implement the literature review process. Table 5 outlines the technologies utilized and the rationale for choosing them. Refer to Appendix A for a glossary of terms.

Table 5: Technologies Incorporated into Solution 1 and their Justification

Required Design Components	Chosen Technologies for Components Incorporation	Justification
Programming/Scripting language	• Python	• Vast variety of libraries that can be imported for key functionalities (Data analysis, API requests, etc.) **(Objective #2)** • Easy to develop LLMs given simplicity of language
Store/Retrieve articles from user's personal library	• Google Cloud	• Platform is highly scalable in its handling of data. ○ Important considering the variable size of personal libraries [9]. (**Objective #2**) • Cloud integrates well with other google technologies utilized. (**Objective #1**)
Retrieve articles from PubMed and Google Scholar	• SerpApi • E-Utilities	• SerpApi is open-source and can be modified to meet the project/user's search criteria [10]. (**Objective #2**) • E-Utilities is PubMed's official API, making it more fast and secure [11]. (**Objective #2, #4, #5**)
Filter articles for literature bias and evaluate/select user-inputted	• IBM Watson NLU	• Platform has a large set of word characteristics analyzed, making it effective for comprehensively identifying bias/objectivity (Appendix A) and

number of articles for review		filtering articles. ● Reported "50% save time" in word analysis [12]. (**Objective #4**)
Produce summaries of each article and a comprehensive literature review	● GPT-4 ● Self-developed LLM	● GPT-4 performs strongly in zero-shot testing (Appendix A), allowing it to generate summaries with a variety of user-inputted criteria [13]. (**Objective #2, #3**) ● Developing a custom LLM for literature review generation by compiling training data from strong human-written reviews will allow for prioritization of important elements. (**Objective #2, #3, #4**)
Produce and compile citations	● Zotero API	● Using Zotero's official API will make citation processing fast and secure [14]. (**Objective #2, #4, #5**)
Produce PowerPoint skeleton	● Google Slides API	● Google Slides' UI is user-friendly and integrates well with other Google technologies [15]. (**Objective #1**) ● Using Google's API for slides will make the PowerPoint generation process fast and secure. (**Objective #4, #5**)
Export summaries, literature review, and citations	● .txt file type	● The .txt file type is easy to read/edit and is universal across computer systems, allowing for copy/paste functionality and text transferability [16]. (**Objective #1, #2**)
Security/Login management	● Google Authenticator 2FA	● Easy to handle, single use 2FA passcode generation **(Objective #1, #5)** ● Integrates well with other Google technologies used **(Objective #1)**

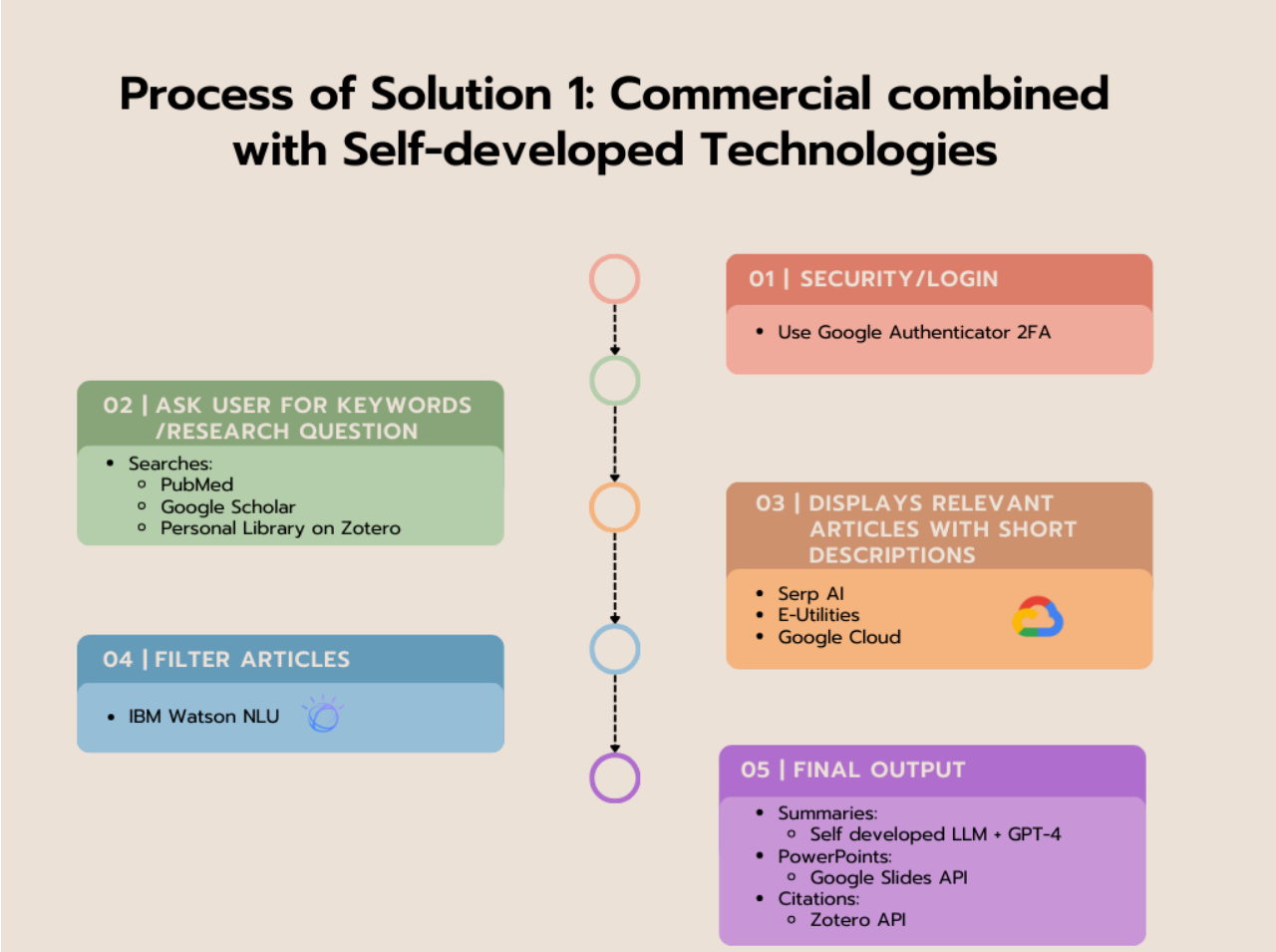


Figure 4. Process of Solution 1.

6.3.2 Solution 2: Low Use of Known Technologies

Solution 2 focuses on manually designing and coding several components of the team's design, and only uses pre-existing security, API's and code for inspiration. It follows the same automated design process demonstrated in Figure 2. The benefit of this solution is its ability to be customizable due to not relying on many external softwares. Table 6 outlines the technologies used and their justification. Refer to Appendix A for a glossary of terms.

Table 6: Technologies Incorporated into Solution 2 and their Justification

Required Design Components	Chosen Technologies for Components Incorporation	Justification
Coding language	• Python	• Lots of libraries and APIs can be accessed through Python [17] which are referenced in the following rows in this

		table. ● Simple programming language that makes it easier to develop a LLM [18] ● Mostly coding ourselves gives the team more control to create something versatile and easy to use. (**Objective #2, #3**)
Retrieve Articles from the Sources	● SerpApi ● E-Utilities API	● Easy to implement into Python. ● Strong service that currently exists.
Summaries Generator	● GitHub made by user miso-belica called sumy	● Able to summarize text by using the most relevant words in the text ● Creator successfully defended their software as their thesis and has over 3500 current users. [20]. ● Been able to reduce word count by 20% ● Contains a list of positively relevant, negatively relevant and relevant words which can easily be changed for clients purposes
Presentations Generator Word Document Generator	● Open XML API	● Zip compression allows the client to store more summaries and PowerPoints. ● Existing exports from Microsoft Word to PowerPoint. ● Secure method of moving documents [21]. (**Objective #5**)
Security	● PyOtp library ● Cryptography library ● Getpass library ● Hashlib library	● Password manager and two-factor authentication (2FA) allows for a more secure software [23].

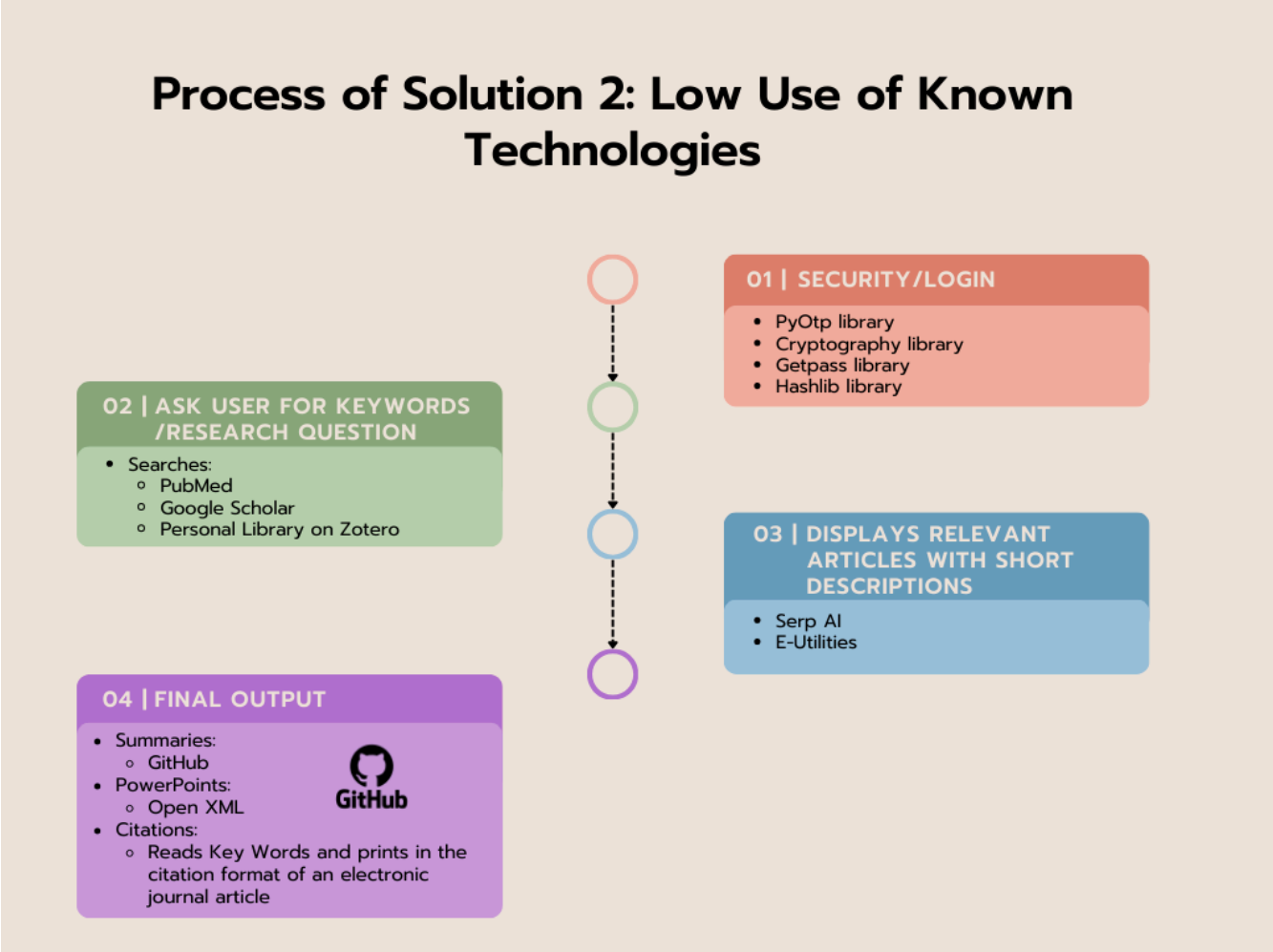


Figure 5. Process of Solution 2.

6.3.3 Solution 3: Complete Combination of Known Technologies

This solution does not require coding for specific components of the design. It is a combination of existing technologies, each performing a certain task, being compatible with Windows OS, and its basic features being free. The following table discusses their implementations. Refer to Appendix A for a glossary of terms.

Table 7: Technologies Incorporated into Solution 3 and their Justification

Required Design Components	Chosen Technologies for Components Incorporation	Justification
Web Scraper (Appendix A)	• Scrapy	• Open source. • Can upload web scraper code onto its cloud [24].

Retrieve Articles from the Sources based on the Keywords	• Perplexity	• Retrieves articles from the source requested by the user (Appendix J).
Generate Summaries	• PopAI	• Presentation generated has an average Flesch-Kincaid Reading Ease of 22.3, calculated through the online Flesch-Kincaid calculator [25] (Appendix J). (**Objective #3**)
Generate Presentations	• Pitch	• Presentation generated has an average Flesch-Kincaid Reading Ease of 6.4 [25] (Appendix J). (**Objective #3**) • Presentations generated can be edited to have black/white background with white/yellow text respectively. (**Objective #2**)
Security	• DUO mobile	• Offers 2FA for additional security [26]. (**Objective #5**) • Users can be monitored by the client. (**Objective #5**)
Produce Citations	• MyBib	• Offers 9,000+ citation styles [27] . • Users can choose citation style. (**Objective #2**)

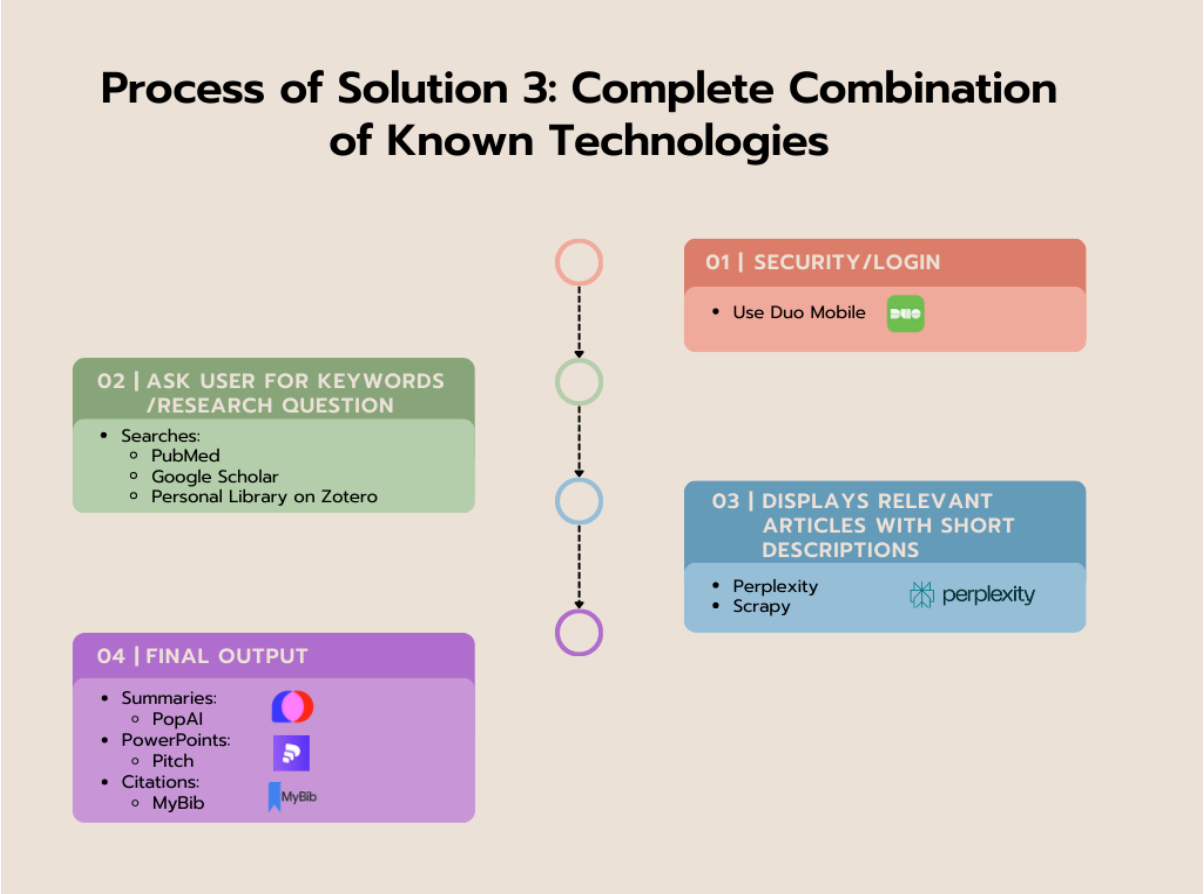


Figure 6. Process of Solution 3.

7.0 Proposed Conceptual Design Specification

The final design selected is Solution 1 - Commercial combined with Self-developed Technologies. The Pugh method was used to determine the proposed solution (Appendix I). In this method, two of the three solutions were compared to Solution 2 in terms of satisfying the five established objectives and were given weighted scores dependent on the ranking of each objective satisfied (Appendix D). Since Solution 1 scored the highest with 5 points (weighted score: 15), it was chosen as the final proposed design.

From consulting the client, there was a strong emphasis on the selection of a few articles from a pool of articles, through selection strategies and inclusion criterias. This process makes the literature review systematic. What differentiates the chosen solution from the other proposed solutions in this manner is the use of IBM's Watson NLU, which allows for both a broader range of inclusion criteria and a more rigorous filtering of articles/papers to find only academically strong/objective ones [12]. Similarly, the use of GPT-4, OpenAI's most advanced LLM [28], produces more accurate summaries of the articles than the other solutions. A self-developed LLM specifically for literature reviews will allow the generated review to be much more rigorous and in-depth as there is control over the training data. In terms of speed and security, while the other solutions utilize third-party APIs and frameworks, the chosen solution uses official APIs wherever possible, thus preserving data integrity and improving data transfer speed. From an

ease-of-use perspective, the chosen solution uses technologies within the Google software ecosystem as opposed to different technologies from different entities, thereby forming a seamless UI and user experience. Thus, the solution's strong satisfaction in the key objectives will produce the outcome that the client desires.

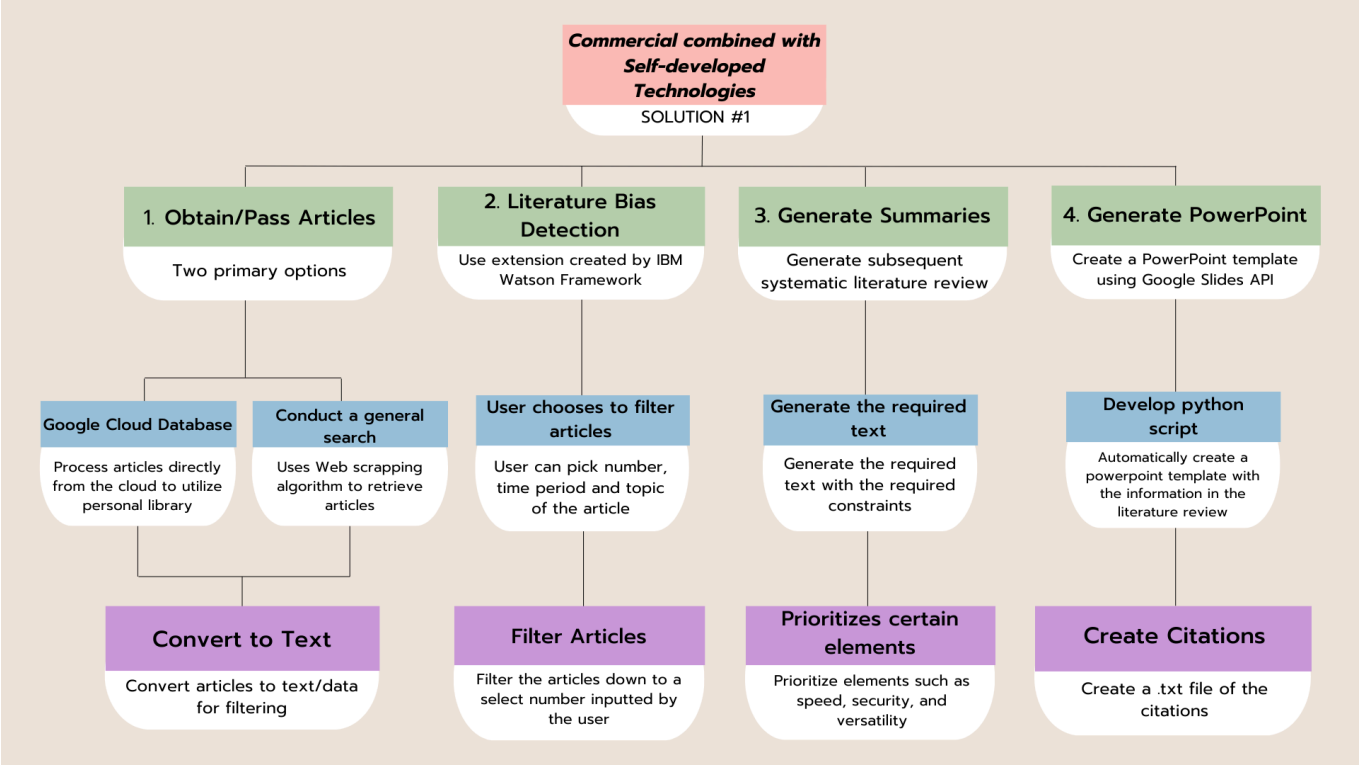


Figure 7. Sequence of the final design.

8.0 Measure of Success

The following section describes how the team will test prototypes against different objectives (Table 3) and determine the successfulness of the test.

Table 8: Testing of Objectives and Determinants of Success

Objective	Test Process	Determinant of Success
Output should span a range of comprehension levels	• Input one article into an AI software used in the solution. • Test output against Flesch-Kincaid Reading Ease.	Summaries score between 60-0 on Flesch-Kincaid Reading Ease.

Solution should be versatile in its input/output (PowerPoint format)	• Input one article into an AI software used in the solution. • Output different PowerPoint templates based on specifications requested by the team.	Outputs PowerPoints of various background colours (black, white) and of various font colours (white, yellow and black)
Solution should be versatile in its input/output (Citation output)	• Input one article into an AI software used in the solution. • Output different citation styles as requested by the team.	Able to cite in IEEE, APA and MLA
Solution should be easy to operate	• Create an interactive slideshow that will represent the UI of the solution. • Get feedback from 3 other teams on the ease of use.	# of clicks for: • data input (**5-8)** • processing information (**2-3**) • obtaining/downloading output (**2-3**) 2+ teams give positive feedback with minimal complaints.

9.0 Conclusion

The first half of the document solely focuses on analysing the engineering problem of not having an autonomous literature review system. The Project Requirements highlight the stakeholders of the project, and its functions, objectives and constraints it must abide by. The second half explores potential solutions to this problem. In summary, the final conceptual design addresses all functions as per the client's objectives and constraints. The input of the system will be keywords or a research question from the user and the program will produce summaries for each article in the form of a Word file. An editable PowerPoint template will also be generated, and citations will be put into a citation manager. Moving forward, the team will carry out the steps outlined in the measures of success (Table 8) when the solution is agreed upon by the client.

10.0 Reference List

[1] PopAi, "PopAi | Your Personal Workplace,". [Online]. Available: https://www.popai.pro

[2] O. AI, "Technical Requirements for Using ChatGPT," *ChatGPTDetector*, Jan. 01, 2024. [Online]. Available: https://chatgptdetector.co/chat-gpt-requirements/

[3] A. Stateczny, K. Gierlowski, and M. Hoeft, "Wireless Local Area Network Technologies as Communication Solutions for Unmanned Surface Vehicles," *Sensors*, vol. 22, no. 2, p. 655, Jan. 2022. [Online]. doi: https://doi.org/10.3390/s22020655.

[4] A.Kassner, "Our Team," Sick Kids. Accessed: Feb. 15,2024. [Online]. Available: https://lab.research.sickkids.ca/kassner/home/people

[5] Natural Sciences and Engineering Research Council of Canada, "NSERC - Discovery Grants (individual) program," Natural Sciences and Engineering Research Council of Canada (NSERC), Jun. 28, 2016. [Online]. Available: https://www.nserc-crsng.gc.ca/ResearchPortal-PortailDeRecherche/Instructions-Instructions/DG-SD_eng.asp

[6] A. Seffah, M. Donyaee, R. B. Kline, and H. K. Padda, "Usability measurement and metrics: A consolidated model," *Software Quality Journal*, vol. 14, no. 2, pp. 159–178, Jun. 2006. [Online]. Available: https://doi.org/10.1007/s11219-006-7600-8

[7] "Guide to Academic Writing Article - Management - University of Canterbury - New Zealand," *web.archive.org*, Jul. 12, 2016. [Online]. Available: https://web.archive.org/web/20160712094308/http://www.mang.canterbury.ac.nz/writing_guide/writing/flesch.shtml

[8] A. Szczepanek, "Password Entropy Calculator," *Omni Calculator*, Feb. 12, 2024. [Online]. Available: https://www.omnicalculator.com/other/password-entropy#faq

[9] A. S. (Year, Month Day). Scaling with Confidence: Leveraging Google Cloud Platform's Scalability Features for Application. Medium. [Online]. Available: https://medium.com/@alex.s12/scaling-with-confidence-leveraging-google-cloud-platforms-scalability-features-for-application-f0a8cb7aae61

[10] SerpApi. (n.d.). Google Scholar API. [Online]. Available: https://serpapi.com/google-scholar-api

[11] Durbin, R., Eddy, S. R., Krogh, A., & Mitchison, G. (1998). Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids. NCBI Bookshelf. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK25501/

[12] IBM. (n.d.). Natural Language Understanding. [Online]. Available: https://www.ibm.com/products/natural-language-understanding

[13] M. Payne and P. Hennis, "Evaluating GPT-4 Zero-Shot Summarization," Width.ai. [Online]. Available: https://www.width.ai/post/evaluating-gpt-4-zero-shot-summarization

[14] Zotero. (n.d.). Web API v3 Basics. [Online]. Available: https://www.zotero.org/support/dev/web_api/v3/basics

[15] E. Magner, "15 Benefits of Using Google Slides," Mvix. [Online]. Available: https://mvix.com/blog/15-benefits-of-using-google-slides/

[16] Lenovo. (n.d.). Glossary: Text File. [Online]. Available: https://www.lenovo.com/ca/en/glossary/text-file/

[17] Python, "The Python Standard Library,". Accessed: Mar.21, 2024. [Online]. Available: https://docs.python.org/3/library/index.html

[18] G. Pai, "Developing Large Language Models - A Simplified Guide," LinkedIn. Accessed: Mar.21, 2024. [Online]. Available: https://www.linkedin.com/pulse/developing-large-language-models-simplified-guide-girish-pai-cckbf/

[19] A.McMurray, "Python is 57x slower than C++ (and 45x worse for the planet)," efinancialcarrers. Accessed: Mar.21, 2024. [Online]. Available: https://www.efinancialcareers-canada.com/news/2023/06/which-programming-language-uses-the-most-energy

[20] Miso-Belica, "Sumy," Github. Accessed: Mar.21, 2024. [Online]. Available: https://github.com/miso-belica/sumy

[21] Microsoft, "Open XML formats and file name extensions,". Accessed: Mar.21, 2024. [Online]. Available: https://support.microsoft.com/en-us/office/open-xml-formats-and-file-name-extensions-5200d93c-3449-4380-8e11-31ef14555b18

[22] M. Abdullahi, "How to Build a Password Manager in Python," The Python Code. Accessed: Mar.21, 2024. [Online]. Available: https://thepythoncode.com/article/build-a-password-manager-in-pytho n

[23] G. Mcgill,"Working with Office "open" XML - just how hard is it?,"Stack Overflow. Accessed: Mar.21, 2024. [Online]. Available: https://stackoverflow.com/questions/9004848/working-with-office-open-xml-just-how-hard-is-it

[24] Scrapy. (n.d.). Scrapy | A Fast and Powerful Scraping and Web Crawling Framework. [Online]. Available: https://scrapy.org/

[25] GoodCalculators. (n.d.). Flesch-Kincaid Calculator. [Online]. Available: https://goodcalculators.com/flesch-kincaid-calculator/

[26] Duo Security. (n.d.). Duo Mobile App: Secure Two-Factor Authentication (2FA) App. [Online]. Available: https://duo.com/product/multi-factor-authentication-mfa/duo-mobile-app

[27] MyBib. (n.d.). Premium Features. [Online]. Available: https://www.mybib.com/premium-features

[28] OpenAI. "GPT-4," openai.com, 2023. https://openai.com/gpt-4

11.0 Appendices

Appendix A: Glossary

The following appendix contains a glossary of terms along with how the team defines them.

Table 9: Glossary of terms

Terms	Meaning
Systematic Literature Review	Methodical process of gathering, evaluating, and summarizing all relevant research on a specific topic to provide a comprehensive overview of the current knowledge in the field
Generative AI	Generative AI describes algorithms that can be used to produce original content such as audio, code, images, text, simulations, and videos. It consists of artificial intelligence, which is the practice of getting machines to mimic human intelligence to perform tasks, and machine learning, which is that they have models that can "learn" from data patterns without human direction.
PubMed	A free search engine by the National Library of Medicine, providing access to biomedical and life sciences literature, including research articles and clinical studies
Zotero	A free software for managing, organizing, and citing research sources.
Journal Clubs	Meeting where members of a particular academic or professional community gather to discuss research articles
Inclusion Criteria	Inclusion criteria are specific requirements used to determine if a participant, study, or item is suitable for inclusion in research, reviews, or analysis.
Search Strategies	Search strategies are systematic approaches used to find relevant information by selecting keywords and employing techniques like Boolean operators and truncation.
API	A software intermediary that allows two applications to transfer data/functionality between each other.

Web Scraper	A web scraper is a software tool or program designed to extract data from websites. It automatically navigates through web pages, retrieves specific content, and then stores that content for further use or analysis.
Open Source	Open source refers to a type of software development model in which the source code of a software program is made freely available and accessible to the public.
Cloud	Refers to a network of remote servers hosted on the internet that store, manage, and process data, as opposed to a local server or personal computer.
2FA (2-factor authentication)	An identity and access management security method that requires two forms of identification to access resources and data.
Large Language Model (LLM)	A sophisticated artificial intelligence system trained on vast amounts of text data. These models are capable of understanding and generating human-like text across a wide range of topics and tasks, including language translation, text summarization, and question answering.
Literature Bias	Literature bias refers to the phenomenon where the available literature on a particular topic is skewed or disproportionately represents certain perspectives, findings, or types of research.
Natural Language Understanding (NLU)	The branch of artificial intelligence and computational linguistics focused on enabling computers to understand, interpret, and respond to human language in a way that is meaningful by analyzing various aspects of language, including syntax, semantics, context, and pragmatics.
Zero-Shot Testing	Zero-shot testing refers to a testing methodology used in natural language processing (NLP) and artificial intelligence (AI) tasks, particularly in language models like GPT (Generative Pre-trained Transformer) models. In zero-shot testing, a model is evaluated on tasks or scenarios for which it has not been explicitly trained. Instead of fine-tuning the model on specific tasks, zero-shot testing assesses the model's ability to generalize and perform adequately on unseen tasks based on its pre-existing knowledge.

User Interface (UI)	UI refers to the visual elements, controls, and interactive components through which users interact with a software application, website, or device.
Github	A platform where creators can create and share code that they have developed.
Python	A High-level programming language that is coded in C.
C	A low-level programming language.
C++	A low-level programming language which is the industry standard for programming.
Open XML	A zip-compressed XML file that carries Microsoft-based files.
Serp-AI	The API that a user would need to use to access PubMed data.
E-Utilities	The API that a user would need to use to access Google Scholar data.
IBM Watson	A data analytic AI that responds to inquiries written through natural language.

Appendix B: Stakeholders

The following appendix describes the stakeholders. Figure 8 compares the interest of the stakeholders against their overall influence to show which stakeholders have the biggest stake in the project.

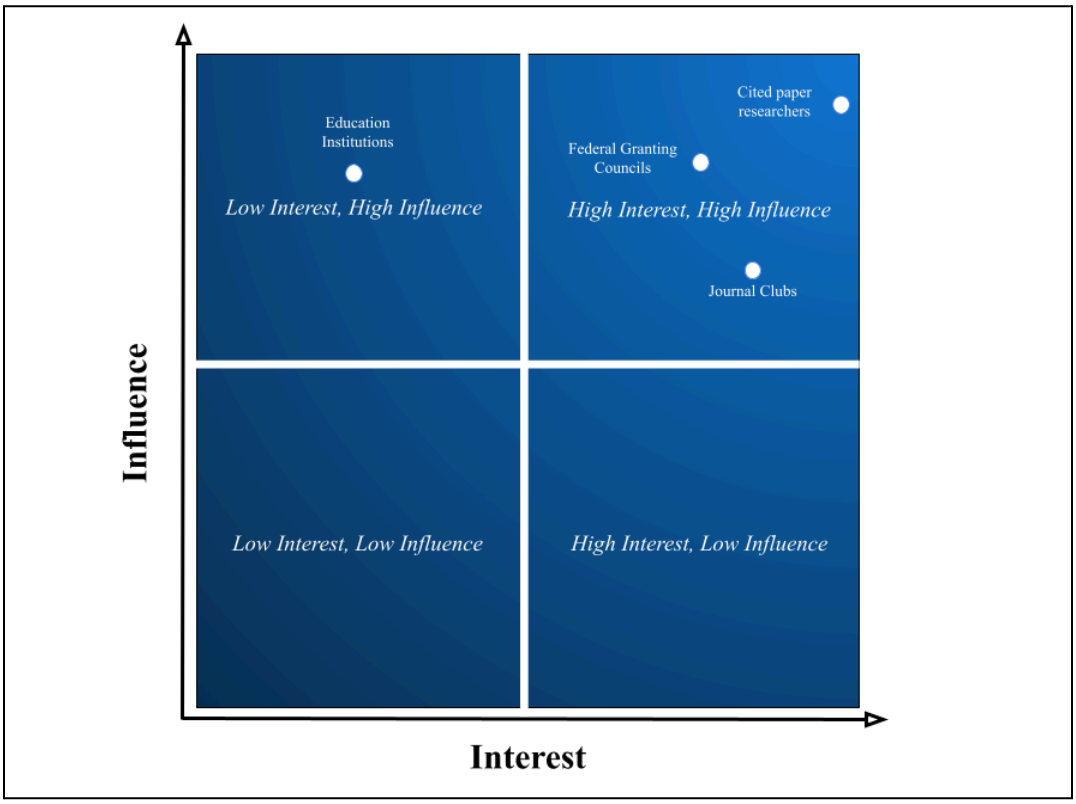


Figure 8. Interest vs. Influence graph to organize stakeholders

Appendix C: Methods for Determining Functions

The following appendix demonstrates the methods used to determine the functions of the design. Figure 9 shows the functional basis to articulate primary and secondary functions. Figure 10 uses the black box method to generate functions from the input and output of mass, energy and information.

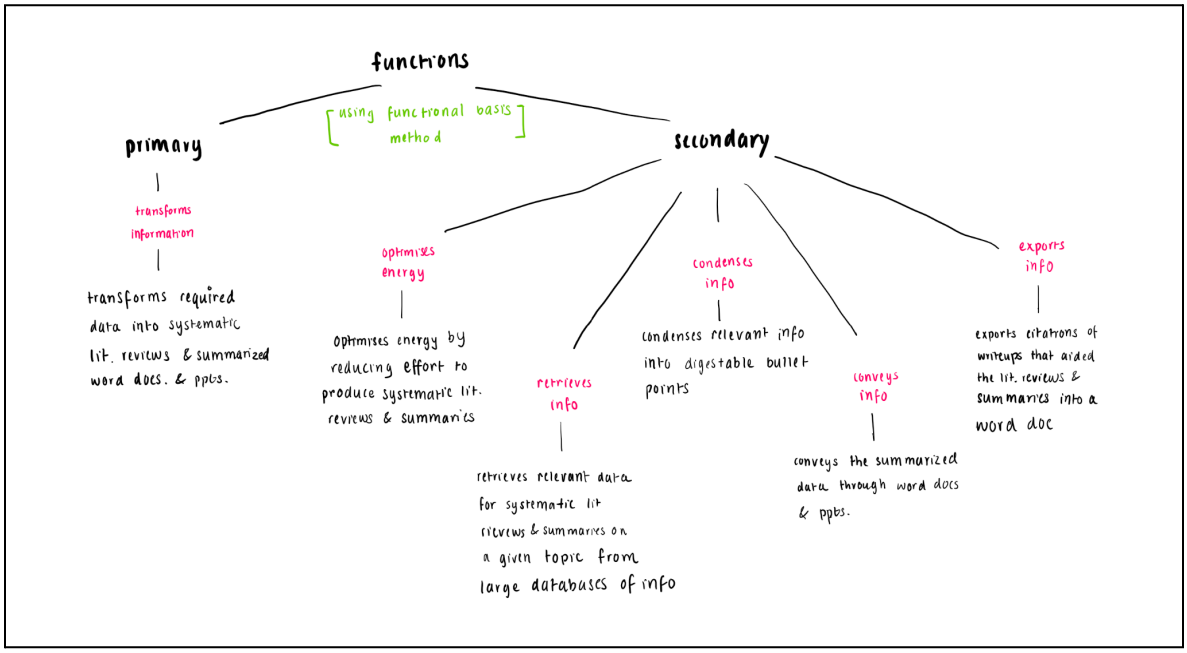


Figure 9. Functional basis to differentiate primary and secondary functions

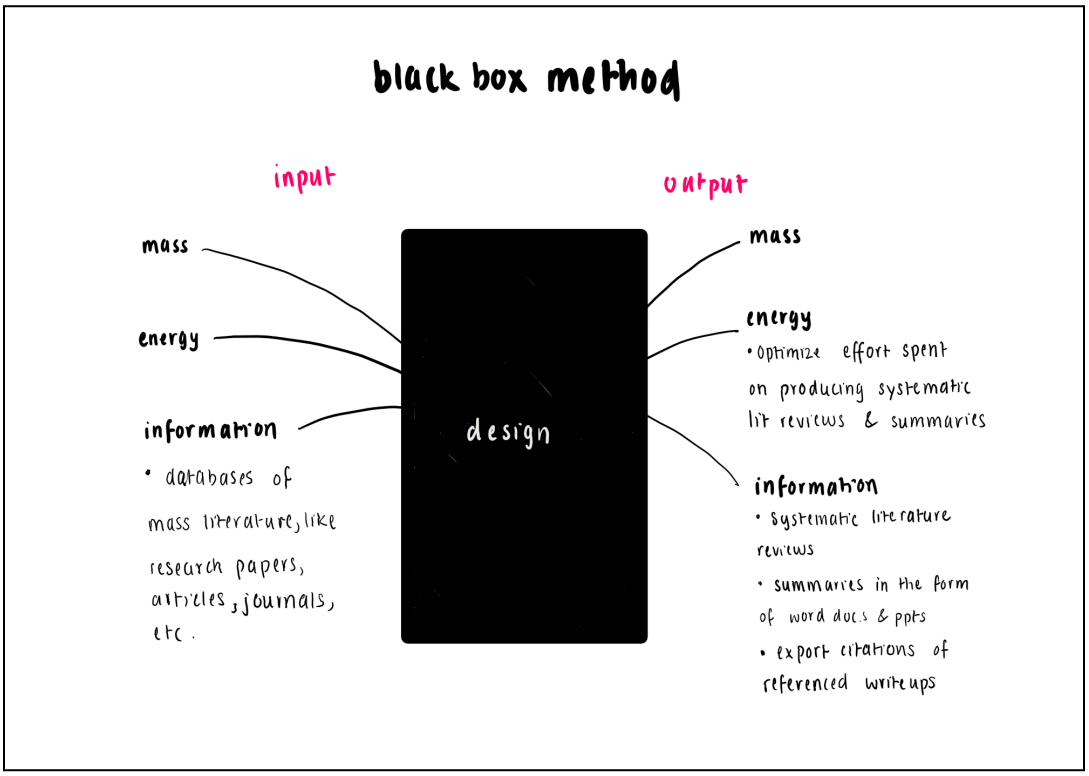


Figure 10. Black-Box method to determine functions

Appendix D: Pairwise Comparison Chart for Objectives

Table 10 performs a pairwise comparison of the objectives to determine which objectives are the most important.

Table 10: Pairwise Comparison of Objectives

	Ease of Operation	**Versatile Input**	**Broad-Ranging Output**	**Fast**	**Secure**	**Accessible**	**Total**
Ease of Operation	X	1	1	1	1	1	5
Versatile Input	0	X	1	1	1	1	4
Broad-Ranging Output	0	0	X	1	1	1	3
Fast	0	0	0	X	0	1	1
Secure	0	0	0	1	X	1	2
Accessible	0	0	0	0	0	X	0

Appendix E: Client Meeting Notes

The following section of the appendix contains the client meeting notes taken by all the members of the team on February 8th, 2024.

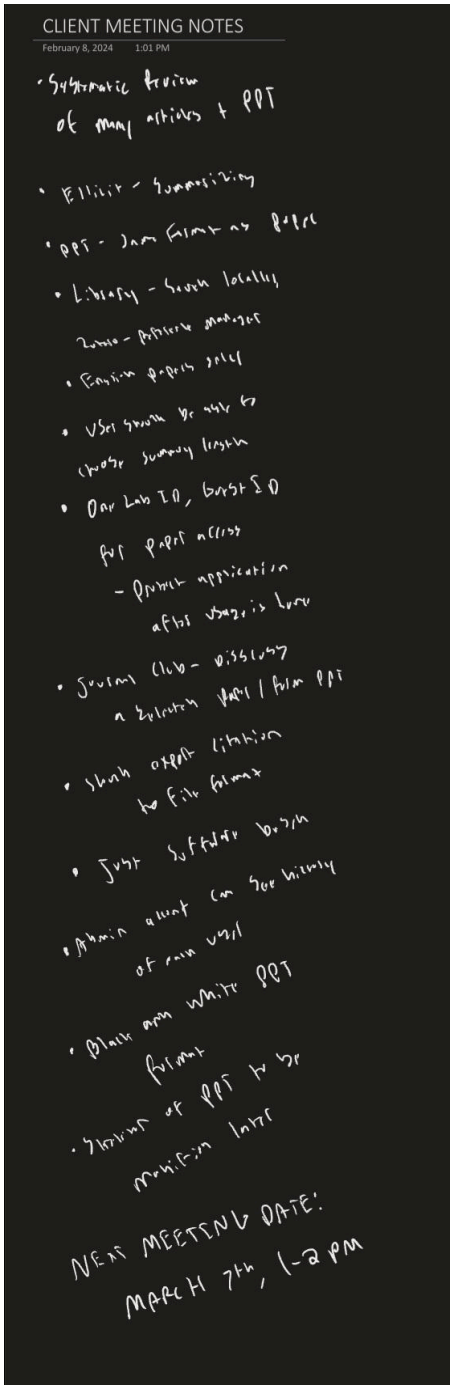


Figure 11. Hisham's client meeting notes.

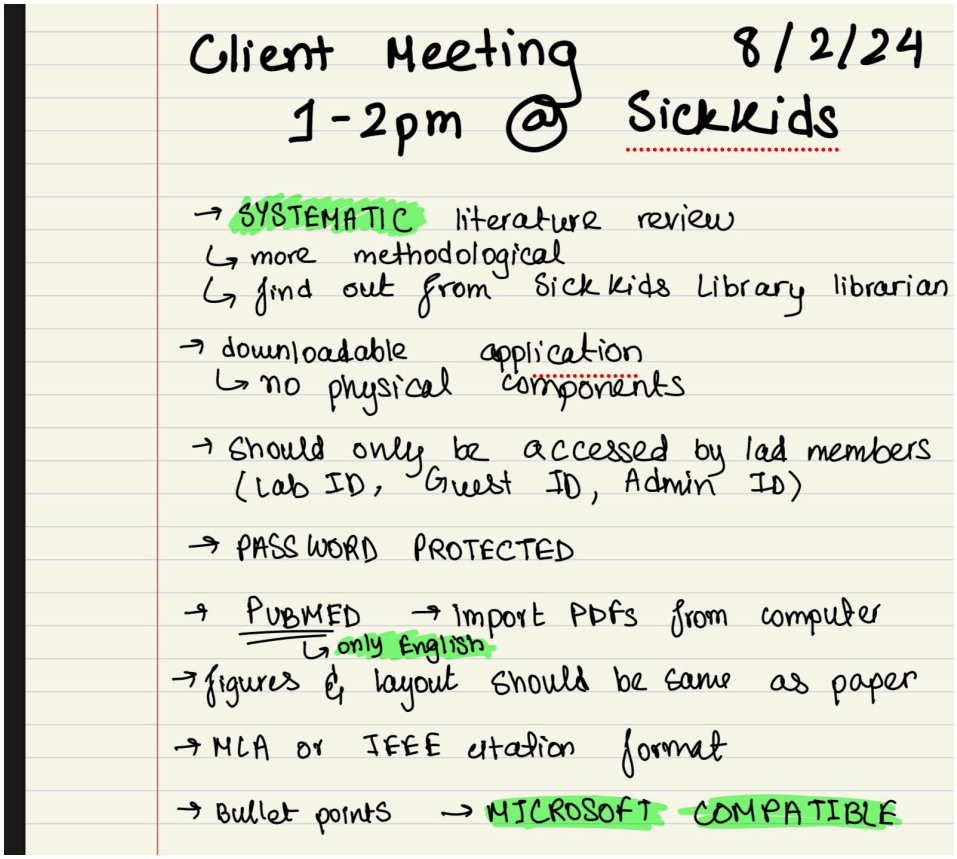


Figure 12. Nameera's client meeting notes.

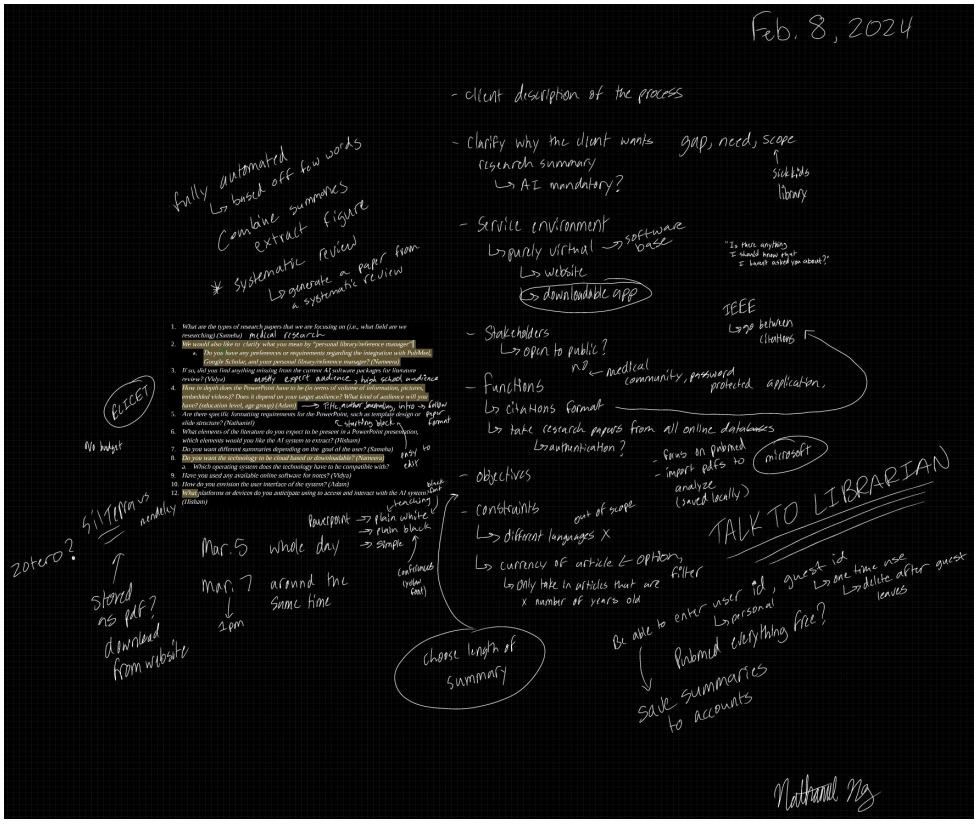


Figure 13. Nathaniel's client meeting notes.

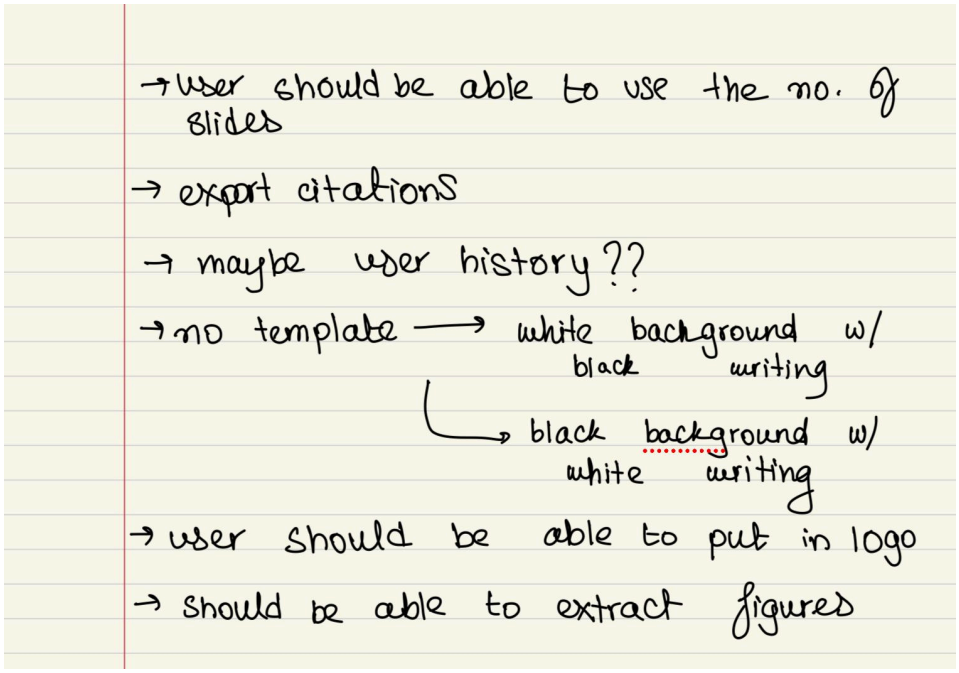
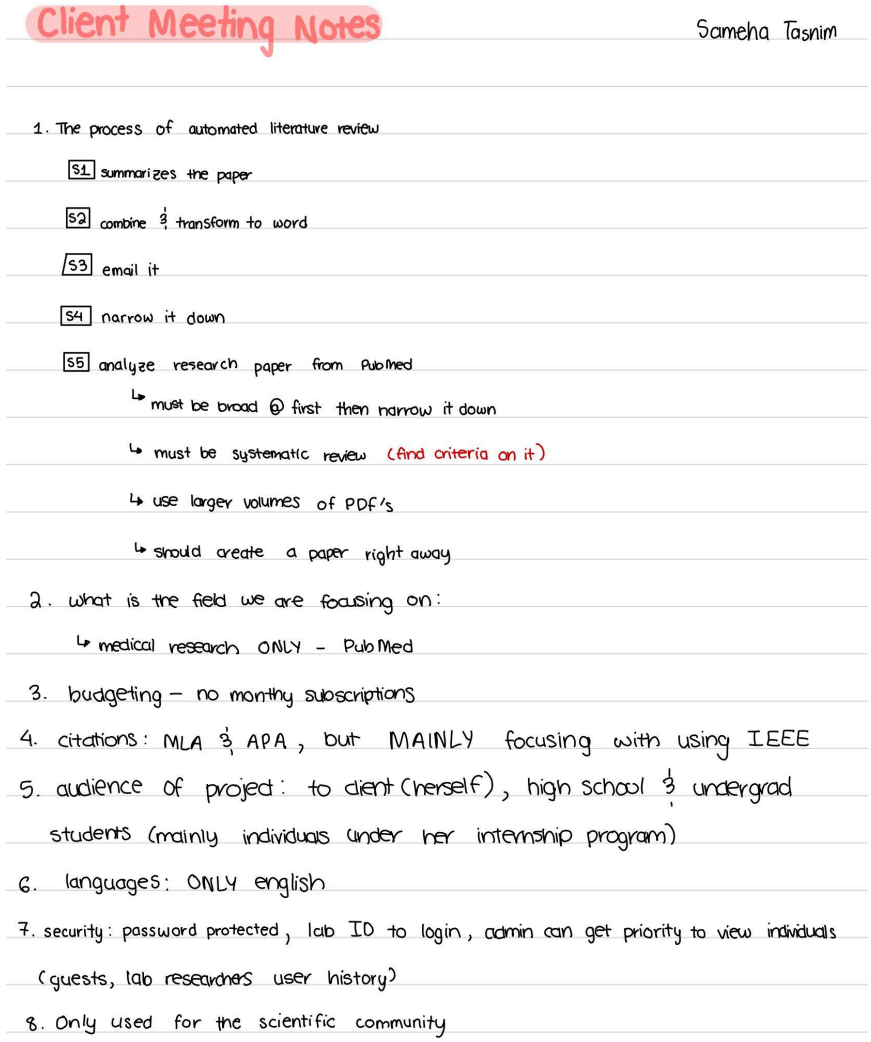


Figure 14. Nameera's client meeting notes.

Figure 15. Vidya's client meeting notes.



Client Meeting Notes — Sameha Tasnim

1. The process of automated literature review
   - S1 summarizes the paper
   - S2 combine & transform to word
   - S3 email it
   - S4 narrow it down
   - S5 analyze research paper from PubMed
     - ↳ must be broad @ first then narrow it down
     - ↳ must be systematic review (find criteria on it)
     - ↳ use larger volumes of PDF's
     - ↳ should create a paper right away
2. what is the field we are focusing on:
   - ↳ medical research ONLY - PubMed
3. budgeting – no monthly subscriptions
4. citations: MLA & APA, but MAINLY focusing with using IEEE
5. audience of project: to client (herself), high school & undergrad students (mainly individuals under her internship program)
6. languages: ONLY english
7. security: password protected, lab ID to login, admin can get priority to view individuals (guests, lab researchers user history)
8. Only used for the scientific community

Figure 16. Sameha's page one of client meeting notes.

Client Meeting Notes (continue)

9. PowerPoint specifications:

- ↳ powerpoint must act as a summary for the paper
- ↳ use microsoft - PowerPoint
- ↳ must have white background with black font colour
  - ↳ OR black background with white OR yellow colour
  - ↳ should include sick kids & UOFT logo

Next Client Meeting: March 7th, 1PM @ sickkids

Figure 17. Sameha's page two of client meeting notes.

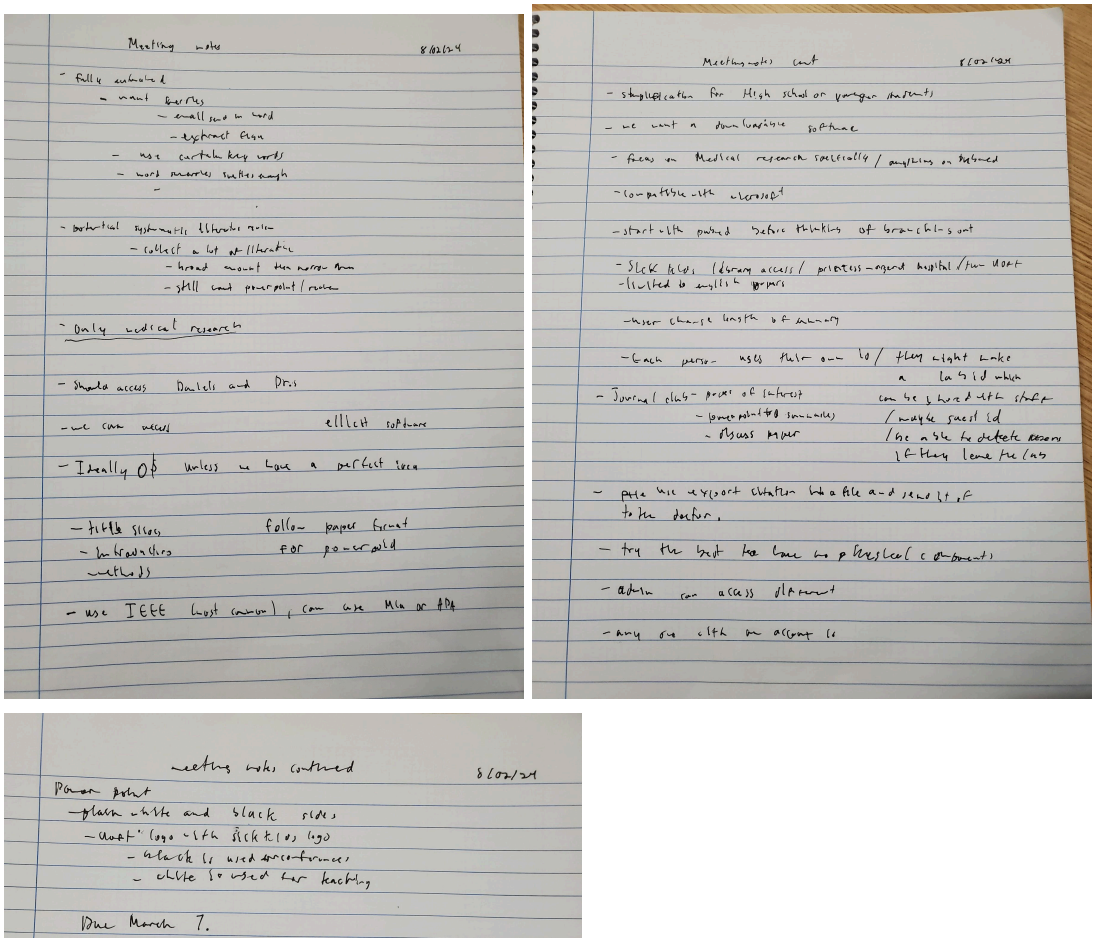


Figure 18. Adam's client meeting notes

The following appendix contains the client meeting notes taken by some of the members on March 4th, 2024.

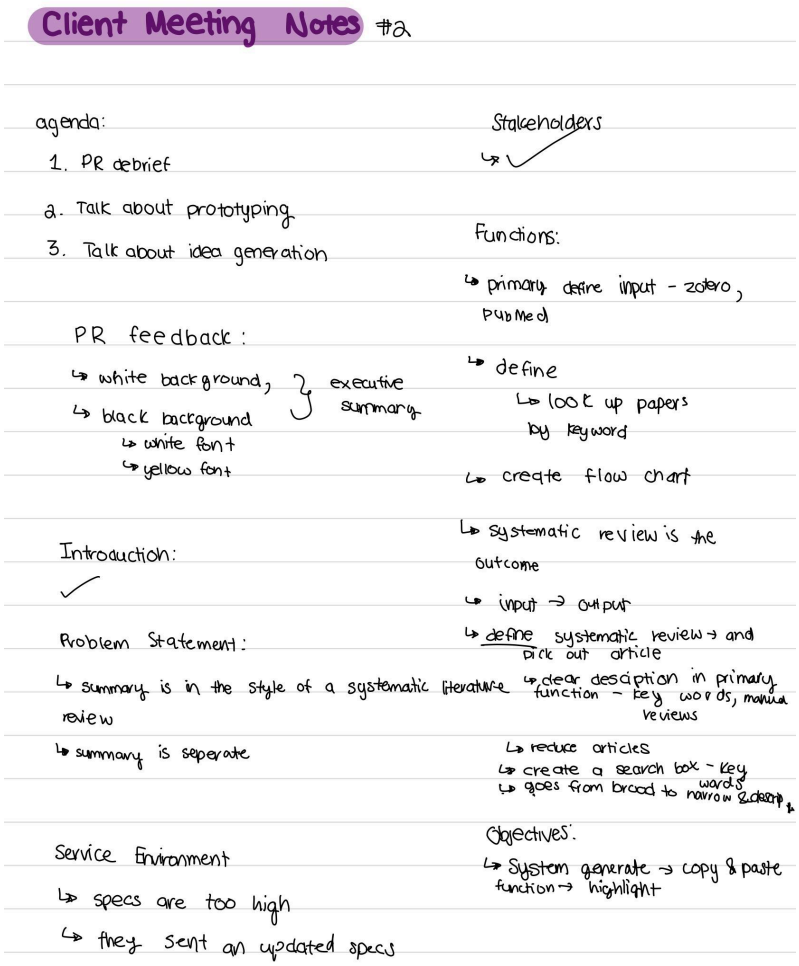


Figure 19. Displays Page 1 of Sameha's Client Meeting Notes.

Figure 20. Displays Page 2 of Sameha's Client Meeting Notes.

Adam's client meeting notes

Elements of PR client found satisfactory

- Gap, need, scope
- Stakeholders: she said everything was fine but it took a while to understand that the section surrounding academic researchers was about **cited** researchers.

Little improvement needed

- Introduction: explain how there will be separate documents for summaries and systematic literature analysis
- Executive summary: specify option between (black background with (white/yellow font)) and (white slide with black font)
- Service environment: lower the hardware specs to accommodate the client equipment, specs were supposed to be emailed to us
- Objectives: include a copy and paste function as another objective
- Constraints: Include Google Scholar, have all documents be editable, and include a flow chart for the flow of data.

Important improvements

- Functions:
    6. Create flowchart to explain process
    7. Define input(does it mean zotero or pubmed etc.)
    8. Same as intro issue
    9. Demonstrate the cutting down process of articles showing that the number of articles goes from "broad to really narrow" based on the summaries
    10. Powerpoints are created based on summaries chosen by the user

Points she made clear with us

- She is fine with the team designing a framework for the project which can include
    5. A logical basis/map
    6. Softwares used
    7. Methods of input
    8. etc..

Figure 21. Displays Adam's Client Meeting Notes.

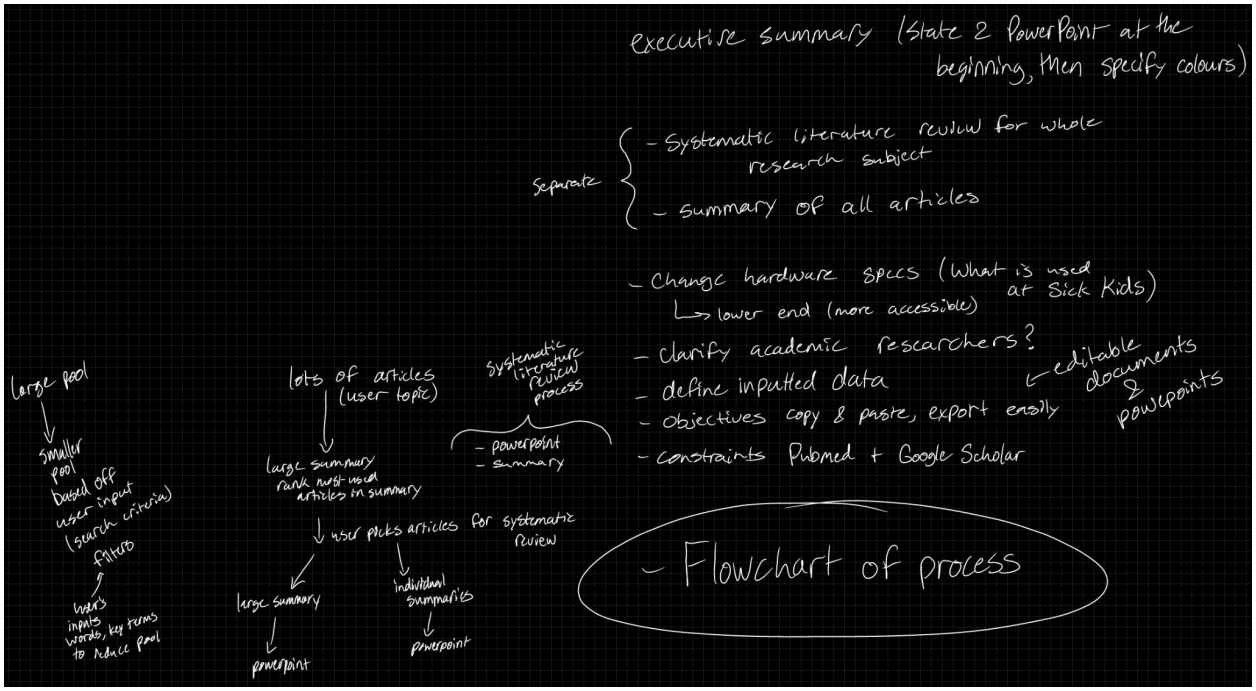
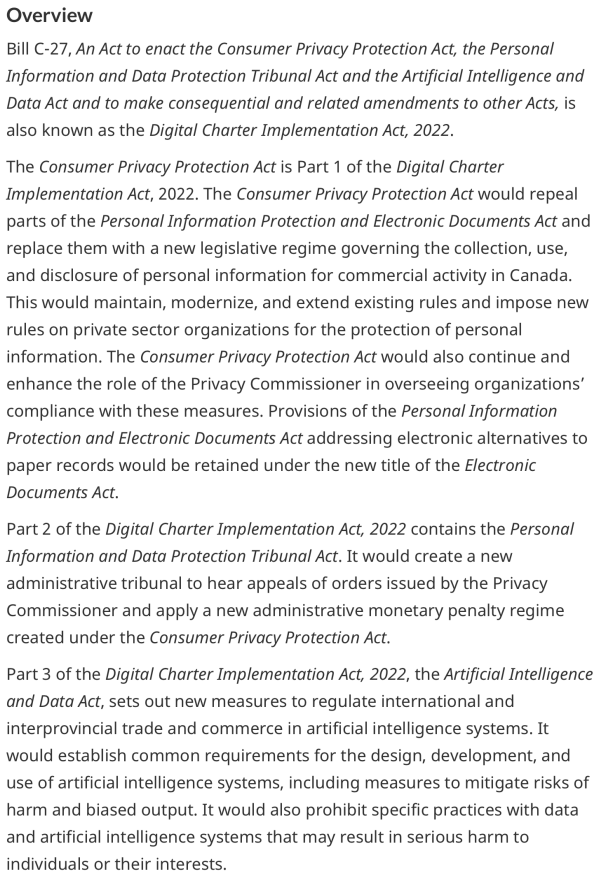


Figure 22. Nathaniel's notes from client meeting 2.

Appendix F: Artificial Intelligence and Data Act (AIDA)

This appendix contains the Artificial Intelligence and Data Act (AIDA) tabled by the Government of Canada in 2022 as a part of Bill C-27.



Overview

Bill C-27, *An Act to enact the Consumer Privacy Protection Act, the Personal Information and Data Protection Tribunal Act and the Artificial Intelligence and Data Act and to make consequential and related amendments to other Acts,* is also known as the *Digital Charter Implementation Act, 2022*.

The *Consumer Privacy Protection Act* is Part 1 of the *Digital Charter Implementation Act*, 2022. The *Consumer Privacy Protection Act* would repeal parts of the *Personal Information Protection and Electronic Documents Act* and replace them with a new legislative regime governing the collection, use, and disclosure of personal information for commercial activity in Canada. This would maintain, modernize, and extend existing rules and impose new rules on private sector organizations for the protection of personal information. The *Consumer Privacy Protection Act* would also continue and enhance the role of the Privacy Commissioner in overseeing organizations' compliance with these measures. Provisions of the *Personal Information Protection and Electronic Documents Act* addressing electronic alternatives to paper records would be retained under the new title of the *Electronic Documents Act*.

Part 2 of the *Digital Charter Implementation Act, 2022* contains the *Personal Information and Data Protection Tribunal Act*. It would create a new administrative tribunal to hear appeals of orders issued by the Privacy Commissioner and apply a new administrative monetary penalty regime created under the *Consumer Privacy Protection Act*.

Part 3 of the *Digital Charter Implementation Act, 2022*, the *Artificial Intelligence and Data Act*, sets out new measures to regulate international and interprovincial trade and commerce in artificial intelligence systems. It would establish common requirements for the design, development, and use of artificial intelligence systems, including measures to mitigate risks of harm and biased output. It would also prohibit specific practices with data and artificial intelligence systems that may result in serious harm to individuals or their interests.

Figure 23. Overview of Artificial Intelligence and Data ACT. Accessed: https://blogs.kent.ac.uk/kbs-library-guide/dissertation/literature-reviews/

Appendix G: Idea Generation, Feasibility Checks and Multivoting

The list and figure below displays the team's process of idea generation which includes methods such as structured brainstorming, eliminating duplicate ideas, feasibility check and the SCAMPER method. The highlighted ideas passed the feasibility check and the multivoting stage which shortened the idea generation list to approximately 50 ideas. Multivoting is represented though the symbol beside each idea where each team member is responsible for 7 votes. The feasibility of the ideas was represented through "FC" on the eliminated ideas meaning that these ideas failed to meet the FOCS. The eliminated ideas are displayed by a strikethrough on the idea.

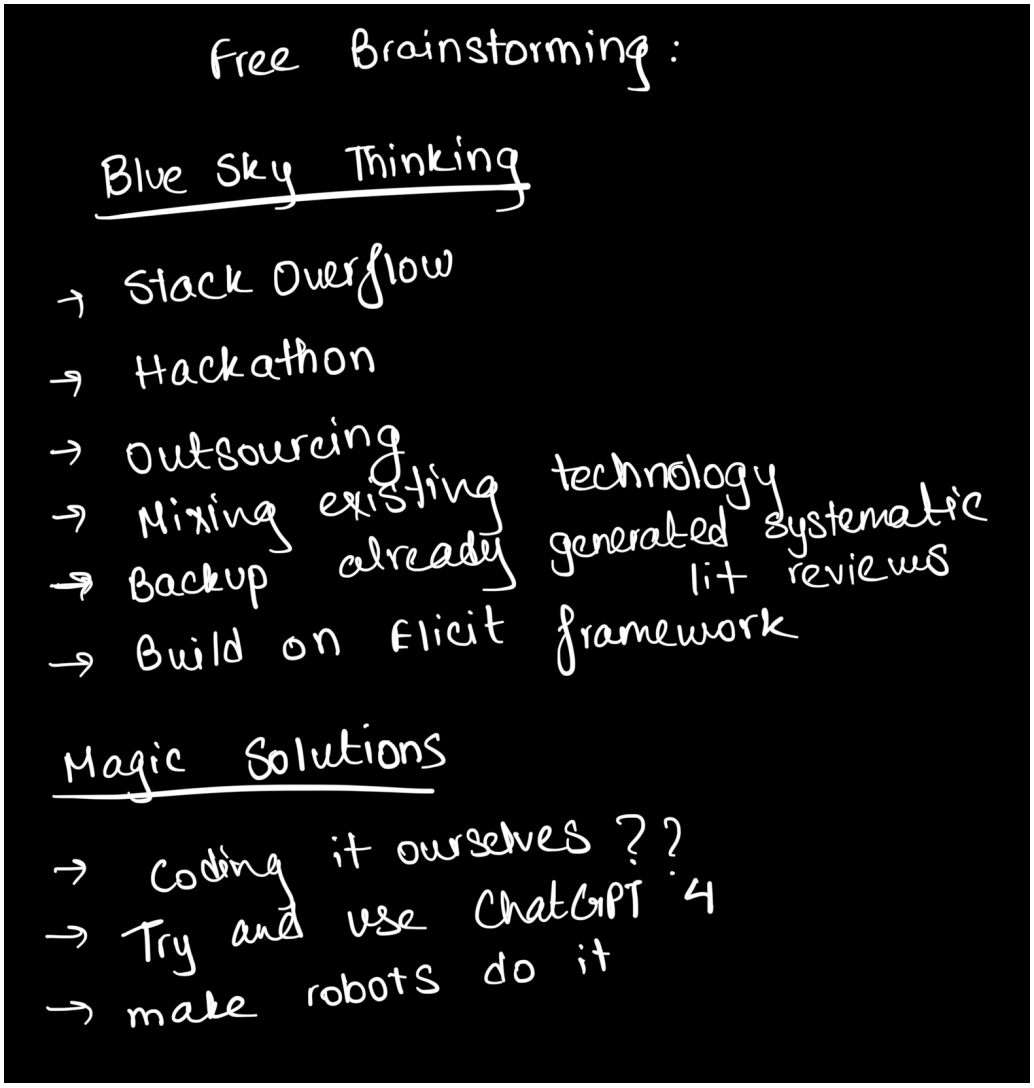


Figure 24. Idea generation process using free brainstorming, blue sky thinking and magic solutions.

Input/Output Versatility:

1. Use PubMed and Zotero APIs 🌐👍🐒🧌
2. Zero shot model
3. Fine-tuned or domain specific model 👍
4. Language representation model
5. Multimodal model (handles both texts and images)
6. Pitch - PowerPoint generator 👍
7. SEMrush AI software that summarizes text
8. Quilbot - has an AI summarizer text section 🐒🧌
9. Web Scraping articles from google scholar 😝🐒🧌
10. Allow user to upload personal articles to an online cloud database 🌐😝🧌
11. Computer Vision Algorithm - used to create the PowerPoint presentation
12. implement a literature bias detector to filter through useable articles 🧌😝

13. ~~Web Scraping - takes data from PubMed in Python~~
14. Incorporate a database for storing citation data, along with APIs for accessing and querying citation information
15. A database for storing metadata and full-text content of research papers, along with a search engine for efficiently querying and retrieving relevant documents.
16. Scribbr🐒
17. MyBib - citation generator
18. Citation machine - citation generator
19. Semantic Search
20. RAKE
21. TextRank
22. YAKE!
23. Genism
24. Spacy
25. SMMRY
26. BART
27. NLTK
28. Sumy
29. Hugging Face Transformers
30. T5 transformer
31. LexRank
32. Pitch
33. Hemmingway Editor
34. Phrasee
35. Copy.ai
36. Content Bot
37. Grammarly
38. Google Slides
39. Clearscope
40. Perplexity
41. GitHub
42. Elasticsearch
43. Algolia
44. Sphinx

Security:

45. Use duo mobile password protection 😝👍🌐
46. Use blockchain for security of system
47. Keycard protection
48. ~~Develop robots that can manually develop systematic lit reviews, summaries, and presentations~~FC
49. Use lastpass (website for free password protection management)

50. Use Bitwarden (website for free password protection management)
51. LLM (Large language model) - used in AI to read inputs as words and reads output as words (will be useful when creating the summary) 😝👍
52. Basic HTML and Hashing Algorithm - for creating passwords
53. Utilize 2FA to make login more secure
54. Encrypted Data
55. Logging and Auditing
56. Behavioral Biometrics
57. OTP Auth
58. Authy
59. LastPass Authenticator
60. Google Authenticator

Ease of Use:

~~61. HTML -for custom site~~ FC
~~62. Implant chips in employees to access computers~~FC
63. Use apple vision pro/VR headset for user to enter information virtually
64. Interactive Visualization Tools 👍

Speed/Efficiency:

65. Compressor.io to optimize page speed 👍
66. PopAI to generate presentation after review generation🐒🌐

Misc.

67. Host a hackathon with $200 cash prize
68. Outsource project to an experienced dev😝🌐🧌🐒
    - Work on UX design and present to dev who codes the solution
69. Hire more people to complete the review in shorter time 🧌
70. Parsing tool - uses AI to read keywords off a website 👍👍
71. Crowdsource data annotation

~~72. Hire students as interns to review journal articles~~FC
~~73. Outsource to India/China~~ FC

74. Create a plug in for web browsers as users browse through PubMed 👍
75. Code a program using python/c to search through keywords 🌐🧌

~~76. Time machine - stops time for surroundings~~ FC
~~77. Start a UofT club~~ FC

78. Use continuous data protection to create copies of literature review/ppt for backup/recovery
79. Backup generated literature review/ppt to aws cloud 👍😝

~~80. Pregabalin~~FC

81. Use account deletion software to remove temporary accounts after term 🌐👍
82. Use Cochrane Database of Systematic Reviews (CDSR) 👍🐒
83. Use Prospero

~~84. Make it a capstone project for final year students~~ FC

~~85. Use stack overflow or piazza (abuse the ego of programmers)~~FC
86. Create a hierarchy system for account authority (admin, user, guest)
~~87. Create a virus (with Kassner) to inject into kids (SickKids) - ages the kids (legally become adults) and speeds up thinking/execution process - virus is cured once lit review is complete~~ FC

Appendix H: Using Morph Charts to Generate Solutions

The following appendix shows the morph chart and how the team used the chart to generate each of our solutions.

Table 11. Morphological chart that displays that each idea has a purpose of meeting the functions.

	Means	Means	Means	Means	Means	Means	Means	Means
Accepting Keywords // Input Versatility	Site APIs	Parsing tools	Semantic Search	RAKE	TextRank	YAKE!	Scrapy	SpaCy
Form summaries	Quillbot	SMMRY	BART	NLTK	Sumy	Hugging Face Transformers	T5 Transformer	LexRank
Form PowerPoint content	Pitch	Hemingway Editor	Phrasee	Copy.ai	ContentBot	Grammarly	Google Slides	Clearscope
Citation generator	Scribbr	MyBib	Citation Machine	Mendeley	Cite This For Me	Citeomatic	OttoBib	EasyBib
Security	DUO mobile	Encrypted Data	Logging and Auditing	Behavioural Biometrics	OTP Auth	Authy	LastPass Authenticator	Google Authenticator
Search Algorithm	Perplexity	GitHub	Elasticsearch	Algolia	Sphinx			

	Means	Means	Means	Means	Means	Means	Means	Means
Accepting Keywords // Input Versatility	Site APIs	Parsing tools	Semantic Search	RAKE	TextRank	YAKE!	Scrapy	SpaCy
Form summaries	Quillbot	SMMRY	BART	NLTK	Sumy	Hugging Face Transformers	T5 Transformer	LexRank
Form powerpoint content	Pitch	Hemingway Editor	Phrasee	Copy.ai	Content Bot	Grammarly	Google Slides	Clearscope
Citation generator	Scribbr	MyBib	Citation Machine	Mendeley	Cite This For Me	Citeomatic	OttoBib	EasyBib
Security	DUO mobile	Encrypted Data	Logging and Auditing	Behavioural Biometrics	OTP Auth	Authy	LastPass Authenticator	Google Authenticator
Search Algorithm	Perplexity	GitHub	Elasticsearch	Algolia	Sphinx			

Figure 25. Demonstration of how ideas in the morph chart were combined.

The following section lists the solutions that each group member came up with along with a short summary of the solution.

Solution 1 - Hisham

- Web Scraping articles from google scholar
- Allow user to upload personal articles to an online cloud database
- Use duo mobile password protection
- LLM (Large language model) - used in AI to read inputs as words and reads output as words (will be useful when creating the summary)
- Implement a literature bias detector (IBM Watson) to filter through useable articles
- Pitch - PowerPoint generator
- Interactive Visualization Tools
- Outsource to experienced dev

- Uses MyBib to create citations in the correct format

Summary of Idea: This idea revolves around two main frameworks: A LLM developed specifically for generating the summaries → systematic literature review and Pitch PowerPoint generator to generate the PowerPoint from the summaries. The user will be able to upload articles from their personal database as well as access articles online through the program's web scraping capabilities, and a literature bias detector extension from IBM Watson will be imported to filter the articles down to a select number inputted by the user. The program will be trained on the developed LLM to generate summaries and a subsequent literature review, and information from the summaries/review will be outsourced to Pitch to create a PowerPoint outline

Solution 2 - Sameha

- Use PubMed and Zotero APIs
- Web Scraping articles from google scholar
- Allow user to upload personal articles to an online cloud database
- Parsing tool - uses AI to read keywords off a website
- Use account deletion software to remove temporary accounts after term
- Backup generated literature review/ppt to aws cloud
- Computer Vision Algorithm - used to create the PowerPoint presentation
- LLM (Large language model) - used in AI to read inputs as words and reads output as words (will be useful when creating the summary)
- Outsource project to an experienced dev

Summary of Idea: This idea does not incorporate any existing generative AI and uses AI techniques to create the automated literature review from scratch. To retrieve the research papers this AI software will use web scraping techniques to attain data from PubMed in python. To minimize the number of articles the user will have the option to search for keywords off the website. This is implemented through the parsing tool. To summarize the research papers the software will use LLM which is used to read inputs as words and reads output as words. Then it would use a computer vision algorithm to create a PowerPoint presentation. Since we do not have these skill sets, we would have to outsource the project to an experienced developer.

Solution 3 - Nathaniel

- Use PubMed and Zotero APIs
- Web Scraping articles from google scholar
- Use duo mobile password protection
- Quilbot - has an AI summarizer text section
- PopAI to generate presentation after review generation
- Use Cochrane Database of Systematic Reviews (CDSR)

Summary of Idea: Use Quillbot and PopAI to generate summary and presentation. Duo Mobile for security. Use PubMed & Zotero APIs and web scraping to help find articles.

Solution 4 - Nameera:

- Use PubMed and Zotero APIs

- LLM (Large language model) - used in AI to read inputs as words and reads output as words (will be useful when creating the summary)
- Parsing tool - uses AI to read keywords off a website

Summary of idea:
We can use LLM and parsing tool so that users can input keywords to find articles from PubMed and can store citations on Zotero

Solution 5 - Vidya

- Use PubMed and Zotero APIs
- Web Scraping articles from google scholar
- Pitch - PowerPoint generator
- Use duo mobile password protection
- PopAI to generate summary & presentation after review generation
- Use account deletion software to remove temporary accounts after term

Summary of Idea: The sources of literature for the systematic literature review are PubMed, Zotero (which can be accessed through APIs - an interface for 2/more computers to communicate with each other) and Google Scholar. This idea combines many existing tools of specific functions: PopAI - to generate the summary, and Pitch - to produce the PowerPoint. For security reasons, DUO mobile will be used for password protection. For additional security, temporary accounts will be removed once not in use by implementing account deletion software.

Solution 6 -Adam

- Use PubMed and Zotero APIs
- PopAI to generate presentation after review generation
- Outsource project to an experienced dev
- Use Cochrane Database of Systematic Reviews (CDSR)
- Use account deletion software to remove temporary accounts after term
- Quill bot with text chunker
- Use our own citation generator.

Summary of Idea: To use the software each user must use the duo authentication to access the software, this access is controlled and can be removed by the owner. Create a program which can input an article into an app called text chunker to later input those chunks into Quillbot due to Quillbot's low word count acceptance. The team will create a program which finds all the necessary elements for a proper IEEE citation and later prints the citation.

Appendix I: Selecting and Comparing the Top 3 Solutions

The following appendix shows the graphical decision chart that was used to choose the top 3 solutions. This appendix also displays how we used the Pugh Method to select the team's final solution. Solution 6 is used as the datum (or standard). The team discussed during an online meeting to figure out the score the other 2 solutions deserve for each objective, in comparison to the datum. The team set the scale of scores to be -3 to 3.

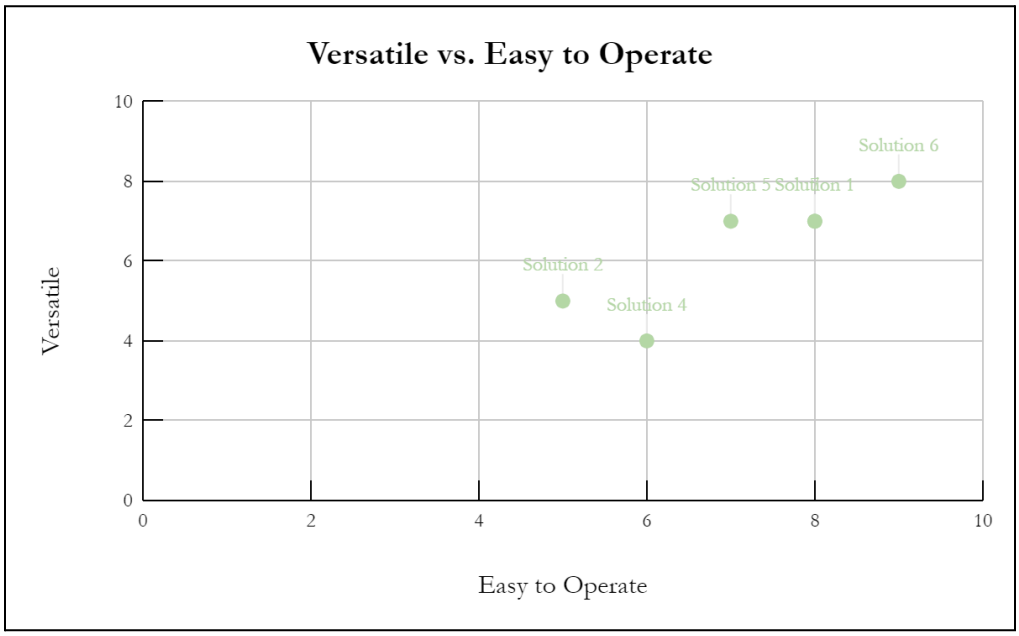


Figure 26. Graphical decision method that gauged versatility against easy to operate.

Table 12: Selecting the final solution using Pugh Method

Objective goal	**Solution 6: Low Use of Known Technologies (Datum) → [new solution 2]**	**Solution 1: Commercial combined with Self-developed Technologies → [new solution 1]**		**Solution 3: Complete Combination of Known Technologies → [new solution 3]**	
		Score	**Justification**	**Score**	**Justification**
Easy to operate (x5 multiplier)	S	1	The idea collaborates software from the Google ecosystem such as 2FA and slides, allowing for seamless integration and ease of use	0	The prototype for this design and the datum have not been made. Therefore, the scores for this objective cannot be compared to one another as of now.
Versatile in its input/output (x4 multiplier)	S	1	The idea allows articles from both personal libraries and online databases to be accessed, and the use of IBM Watson NLU for processing and filtering articles allows	-1	Choices can be made for the visuals of presentations. However, this process cannot be done with a single prompt. Edits need to be made to the presentation to get the desired visuals.

					Other than that, there are also 9000+ choices for citation styles.
Output should span a broad range of comprehensive levels (x3 multiplier)	S	1	Use of OpenAI's powerful GPT4 LLM will allow for more variety of comprehension in the outputted summaries, and the training data set for the custom LLM will incorporate articles of all mentioned Flesch-Kincaid levels	-1	Output produced is within the objective goal metric. However, it is consistently on the lower scale (less than 30). Summaries have the reading level of 'Very difficult to read'. In comparison, the datum is only limited in its comprehension levels.
Fast to use (x1 multiplier)	S	1	Utilization of official APIs for different aspects of the idea instead of 3rd party APIs, as well as using the Google ecosystem of software and IBM Watson NLU, allows the program to run faster than self-developing the aspects that consume the most time (article retrieval, text generation, PowerPoint generation)	-1	Every aspect of this design incorporates an existing technology. Each step is dependent on the other and can only start once its predecessor step is complete. So, this design is likely slower than the datum.
Secure (x2 multiplier)	S	1	The solution incorporates the same security features as the datum in terms of login. However, using official APIs allows for data integrity and protection - making the literature review process more secure	0	The solution incorporates the same security features as the datum.

The final scores are:
Solution 1: 5 (weighted 15)
Solution 5: -3 (weighted -8)
[Solution 6 as datum has score 0]

Since solution 1 has the largest score, it is the final solution.

Appendix J: Components and Testing of Solution 3

The appendix shows the steps conducted as evidence to acquire the numerical values listed in section 6.3.3.
The following steps were carried out:

1. Prompt Perplexity to provide the names of articles based on given keyword/s
2. Prompt PopAI to generate a summary on one of the articles given by Perplexity
3. Prompt Pitch to generate a presentation based on the same keyword/s given to Perplexity
4. Calculate the Flesch-Kincaid Reading Ease of the text in the summary and the presentation using the online calculator

These steps were conducted twice:

1. First time with the keyword: pediatrics (topic 1)
2.

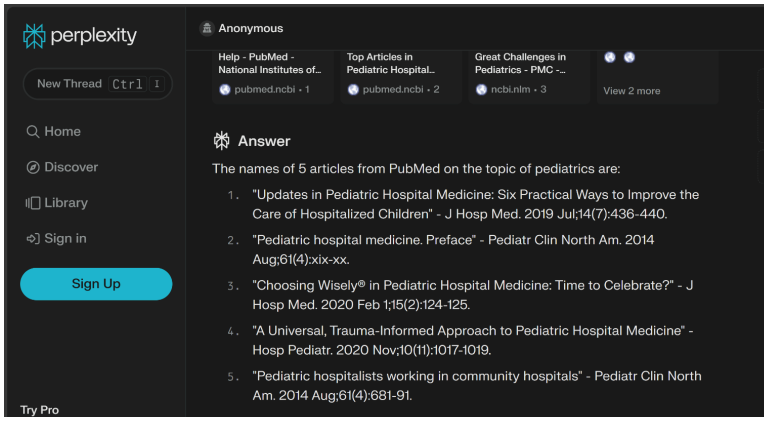


Figure 27. Names of articles generated by Perplexity for topic 1.

The following is the summary generated by PopAI for topic 1.
The document "Updates in Pediatric Hospital Medicine: Six Practical Ways to Improve the Care of Hospitalized Children" presents a review of six significant articles published in 2018, focusing on pediatric hospital medicine. The authors selected these articles based on their quality and applicability to changing practices in the hospital setting or prompting further research. The key findings from the selected articles include the impact of interventions on improving compliance with bronchiolitis guidelines, the development of a calculator to improve testing for urinary tract infections in young children, the underappreciated nonmedical costs of hospitalizations, the effectiveness of a progress note template in electronic health records, the time to pathogen detection for febrile infants, and the development of a high-value care tool for family-centered rounds. The review provides a comprehensive overview of each study, including background, study methods, results, limitations, important findings, and implications for practice and future research.

The review process involved the selection of 18 English-language journals across relevant domains, with a median Hirsch index of 131, to identify articles with novel hypotheses, significant results, and the potential to change pediatric hospital medicine practices. The selected articles were presented at the 2019 Society for Hospital Medicine national conference, where the authors highlighted the most impactful studies. Each selected article was thoroughly examined, providing detailed insights into the study background, results, limitations, and implications for practice and future research. The findings offer valuable contributions to pediatric hospital medicine, addressing various aspects of care and highlighting the need for interventions to improve compliance, reduce unnecessary testing and treatment, quantify nonmedical costs, enhance documentation quality, and promote high-value care in family-centered rounds.

The document also discusses the limitations of each study, such as the inability to determine the individual or subset of interventions responsible for changing provider behavior, potential bias in coverage due to the selection of journals, and the need for further validation of tools and calculators in different settings. The implications of the findings emphasize the importance of reinforcing guideline implementation, guiding

testing for urinary tract infections, supporting families facing financial and social hardships, combatting note bloat in electronic health records, and measuring the frequency and content of high-value care teaching in pediatric inpatient settings. Overall, the review offers a comprehensive and insightful analysis of the selected articles, providing valuable information for healthcare professionals involved in pediatric hospital medicine.

In conclusion, the review on "Updates in Pediatric Hospital Medicine" provides a detailed examination of six impactful articles, offering a thorough understanding of their significance in improving the care of hospitalized children. The comprehensive analysis of each study, along with their limitations and implications, contributes to the advancement of pediatric hospital medicine practices and highlights the potential for further research and interventions in this field.

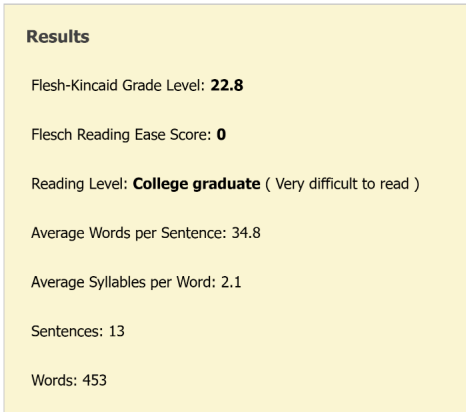


Figure 28. All scores of the summary of topic 1 on the online Flesch Kincaid calculator

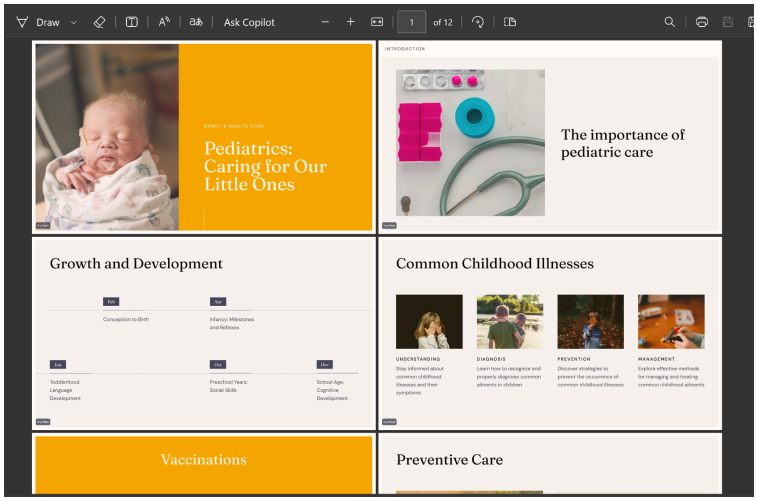


Figure 29. Presentation generated for topic 1

Flesh-Kincaid Grade Level: **7.6**

Flesch Reading Ease Score: **44**

Reading Level: **College** (Difficult to read)

Average Words per Sentence: 2.1

Average Syllables per Word: 1.9

Sentences: 148

Words: 314

Figure 30. All scores of the presentation of topic 1 on the online Flesch-Kincaid calculator

3. Second time with the keywords: pediatrics, adolescents

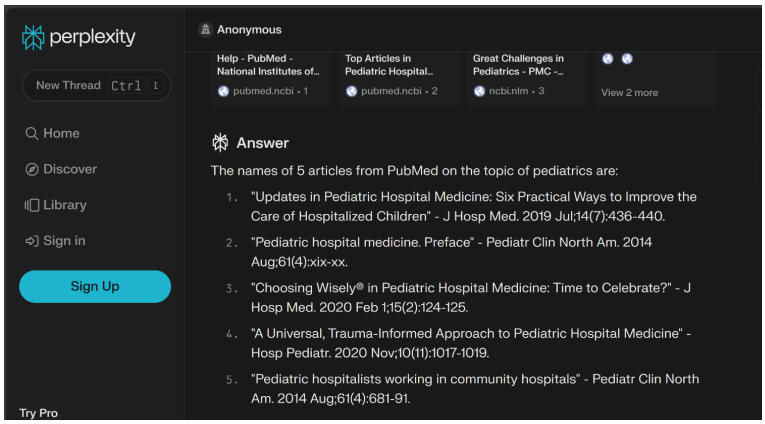


Figure 31. Names of articles generated by Perplexity for topic 2

The following is the summary generated by PopAI for topic 2.

The editorial in the Journal of Adolescent Health (JAH) delves into the multifaceted aspects of pubertal development in adolescents, presenting five articles that explore various dimensions of this critical developmental milestone. The studies investigate pubertal characteristics in diverse populations, emphasizing the impact of early puberty on emotional and behavioral health, as well as the influence of environmental stressors, such as poverty, on pubertal onset. The findings suggest a stabilization of the age of pubertal onset over the past decade in primarily White cohorts, contrasting with earlier reports indicating a decrease in the age of pubertal onset over the past 50 years, particularly in girls. The need for normative data that encompass diverse racial, ethnic, and weight spectrums is emphasized, as these factors significantly influence pubertal timing and associated health outcomes.

Furthermore, the editorial highlights the association between early pubertal development and emotional and behavioral problems, shedding light on the potential vulnerability associated with early environmental adversity. The impact of poverty on internalizing symptoms in adolescents is also explored, suggesting a correlation between neighborhood poverty, early puberty timing, and chronic stress. While the articles touch on the psychological impact of early puberty, a gap in understanding the effects of puberty on physical health outcomes is noted, calling for further exploration of the long-term implications of variations in pubertal onset, including the potential links to physical health issues such as short stature and cardiometabolic diseases. Additionally, the limitations of using self-report assessments to evaluate pubertal development are acknowledged, emphasizing the need for validated self-report staging tools to optimize accuracy in future studies.

The editorial underscores the critical need for up-to-date and relevant data to redefine "precocious puberty," particularly as traditional definitions may be outdated, especially in light of potential shifts in pubertal timing over time. The lack of diverse normative data raises concerns about the potential misclassification, misdiagnosis, and mismanagement of adolescents, particularly if standards are based on primarily White populations. The call for understanding triggers of early puberty and contemporary pubertal trends, as well as the need to consider a broad racial and ethnic

spectrum in research endeavors, underscores the complexity and significance of this developmental stage. The editorial concludes by emphasizing the necessity of continued research to fill the existing knowledge gaps in understanding pubertal development and its implications for adolescent health.

In summary, the editorial in the Journal of Adolescent Health provides a comprehensive overview of the current understanding of pubertal development in adolescents, highlighting the need for diverse normative data, the impact of early puberty on emotional and physical health, and the limitations of self-report assessments. It underscores the importance of ongoing research to better understand the complex interplay of factors influencing pubertal timing and its implications for adolescent health.

Results

Flesh-Kincaid Grade Level: **21.8**

Flesch Reading Ease Score: **0**

Reading Level: **College graduate** (Very difficult to read)

Average Words per Sentence: 32.3

Average Syllables per Word: 2.1

Sentences: 14

Words: 452

Figure 32. All scores of the summary of topic 2 on the online Flesch Kincaid calculator

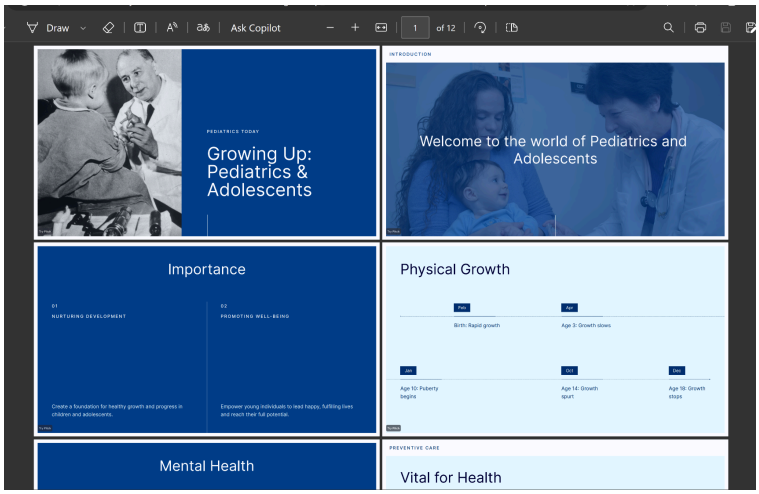


Figure 33. Presentation generated for topic 2

Results

Flesh-Kincaid Grade Level: **5.2**

Flesch Reading Ease Score: **61.1**

Reading Level: **8th & 9th grade** (Plain English)

Average Words per Sentence: 1.9

Average Syllables per Word: 1.7

Sentences: 143

Words: 265

Figure 34. All scores of the presentation of topic 2 on the online Flesch-Kincaid calculator

The averages of the Flesch Kincaid Reading Ease of the two scenarios were taken for the summaries and presentations. These averages are listed in 6.3.3.
The following appendix also contains other visuals of how some components of solution 3 look like.

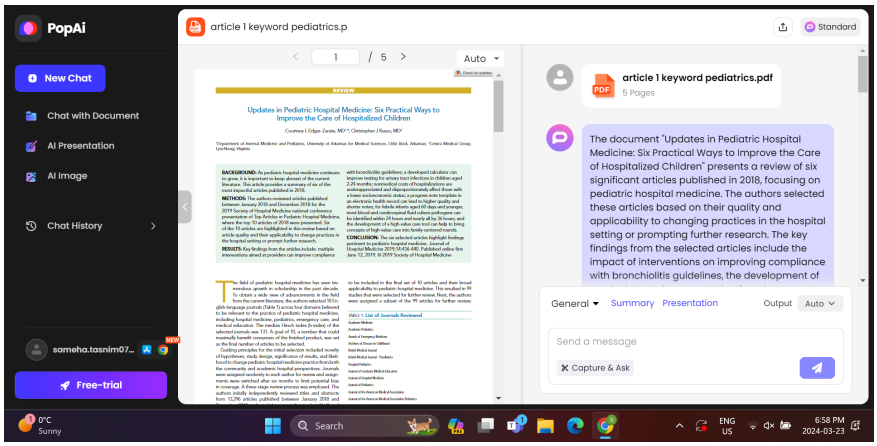


Figure 35. Workspace of PopAI

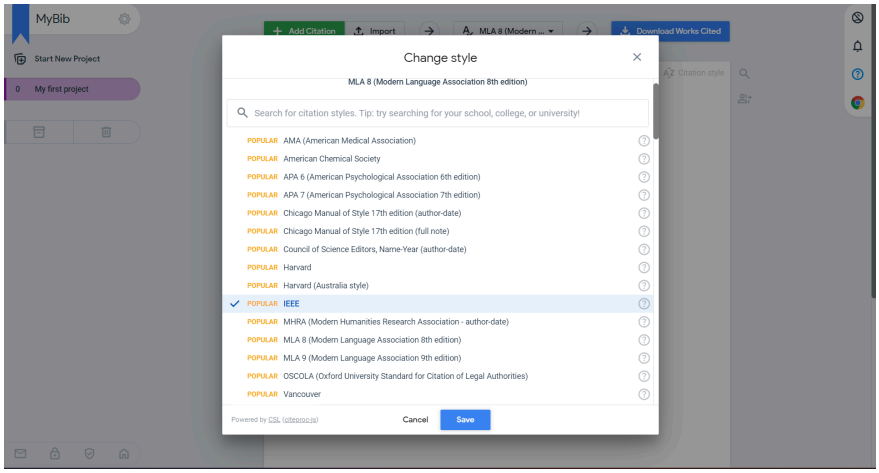


Figure 36. Different Citation Styles offered by MyBib

Appendix I: Attribution Table

Tutorial #:	0119	Team #:	030
Assignment:	CDS	Date:	03/24/2024

The Attribution Table is a major resource used by your TA in determining whether there was equal contribution to the team assignment. If your TA determines that there was significant under contribution, then they may apply an individual penalty to the under contributing team members' grade. As a future professional engineer you should NOT sign any document you have not read and do not agree with.

The Attribution Table must be completed, signed by all team members, and included as an appendix of your assignment AND uploaded to your MS Teams team channel. Teams who do not submit a completed form, including those that submit an incomplete form, such as one missing a team member's signature, will receive zero on the assignment. The team may submit a petition to the ESP Office if they feel the lack of signature is through no fault of the team.

The Attribution Table should accurately reflect each team members' contribution to the document. Be sure to keep a copy of this form for the team's records.

If there are irreconcilable differences that are preventing all team members from signing the attribution table then each team member must write a letter (<one page) explaining their position on the difference and suggest a solution. These letters much be submitted to the TA.

As with any engineering statement this attribution table must be backed by credible evidence. In most cases this will be found either in the Google Docs document revision history, or your engineering notebook. Making fraudulent claims in an Attribution Table displays intent to deceive and is a serious academic offence.

Section	Student Names					
	Adam	Hisham	Vidya	Nameera	Nathaniel	Sameha
1.0 Introduction	ET	ET	ET, OR1, RS1	ET,	ET	
2.0 Problem Statement	ET	ET			WD, ET	
2.1 Gap				WD,	WD, ET	
2.2 Scope					WD, ET	
2.3 Need				WD,	WD, ET	

3.0 Service Environment					ET	WD, ET, RS1, MR
4.0 Stakeholders	WD,MR,ET,OR1,				ET	ET, OR1, OR2
5.0 Detailed Requirements		ET			ET	
5.1 Functions		ET	WD, MR, ET, OR1		ET	
5.1.1 Primary Functions			WD, MR, ET, OR1		ET	
5.1.2 Secondary Functions			WD, MR, ET, OR1	ET	ET	OR2
5.2 Objectives		WD, MR, ET, RS2, OR3			ET	
5.3 Constraints				RS, WD, ET	ET	MR, ET
6.0 Generation, Selection and Description of Alternative Designs		MR, ET		WD, ET, MR	ET	WD, ET, MR
6.1 Idea Generation Process		MR, ET		WD, ET	ET	OR2, ET
6.2 Alternative Design Selection Process		MR, ET	OR3		ET	WD, ET, MR,

6.3.0 Alternative Design Descriptions		ET			ET	
6.3.1 Solution 1		WD, MR, ET, OR3, RS2, RS9		OR1,	ET	OR2, ET
6.3.2 Solution 2	WD, ET, MR,OR1,RS9,RS 10,RS11,RS12	ET		OR1	ET	OR2, ET
6.3.3 Solution 3		ET	RS2, RS3, RS4, RS5, RS6, RS7, RS8, WD, MR, ET, OR3, OR1	OR1	ET	OR2, ET
7.0 Proposed Conceptual Design Specification		MR, ET		WD	ET	ET, MR
8.0 Measures of Success		ET			WD, ET	
9.0 Conclusion			WD	WD,	ET	ET, MR
10.0 Reference List					WD, ET	
11.0 Appendices	OR3, OR1,WD,ET	WD, MR, ET, OR3	RS2, RS4, RS5, RS6, RS8, WD, MR, ET, O1, O2	ET, RS,	WD, ET	WD, ET, MR, O1

Fill in abbreviations for roles for each of the required content elements using the abbreviations found on the next page. You do not have to fill in every cell.

RS – Research (give details below) WD – Wrote Draft MR – Major Revision ET – Edited	FP – Final Proofread of COMPLETE DOCUMENT verifying for flow and consistency OR – Other (give details below)

If you put RS (research) please add a number identifier such as RS1, RS2, etc. Give the research question / topic:

RS1: Literature Review vs Systematic Literature Review

RS2: Flesch-Kincaid Grade Calculator

RS3: Scrappy and its features

RS4: Perplexity and its features

RS5: PopAI and its features

RS6: Pitch and its features

RS7: DUO mobile and its features

RS8: Number of Citation Styles offered by MyBib

RS9 Google Scholar and PubMed API's

RS10 Benefits and Capabilities of using Python

RS11 Available Python libraries

RS12 Preexisting programs in GIthub

If you put OR (other) please add a number identifier such as OR1, OR2, etc. Explain the role below:

OR1: Made diagrams

OR2: Created digital diagrams

OR3: Made tables

By typing your name below to sign, you verify that you have:

- Read the attribution table and agree that it accurately reflects your contribution to the associated document.
- Written the sections of the document attributed to you and that they are entirely original.
- Accurately cited and referenced any ideas or expressions of ideas taken from other sources according to the standard specified by this course.
- Read the University of Toronto Code of Behaviour on Academic Matters and understand the definition of academic offense includes (but is not limited to) all forms of plagiarism. Additionally, you understand that if you provide another student with any part of your own or your team's work, for whatever reason, and the student having received the work uses it for the purposes of committing an academic offence, then you are considered an equal party in the offence and will be subject to academic sanctions.

Student #1 Name Sameha Tasnim Student #5 Name Adam Abreha

Student #2 Name	Nameera Ahmed	Student #6 Name	Hisham Kaleem
Student #3 Name	Vidya Mukherjee	Student #7 Name	
Student #4 Name	Nathaniel Ng		