1. **What are the names and NetIDs of all your team members? Who is the captain? The captain will have more administrative duties than team members.**

Our team consists of Ryan Wolf (ryantw2) who will serve as our team captain, Neal Tinaikar (nealt2), Meghan Lu (meghan7), and Eileen Yang (eileeny2).

2. **What topic have you chosen? Why is it a problem? How does it relate to the theme and to the class?**

We have chosen Intelligent Browsing to integrate intelligence into the web browser experience. As the interface to the Internet, web browsers give users the power to search for and access specific sites and pages that would otherwise be lost in the boundless web space. If the user finds a site especially useful, they can save, or "bookmark", it for easy access from the browser. However, currently there is no easy way to search through the bookmarked pages for keywords or phrases if the website name or user-created bookmark name is forgotten. We seek to provide a solution to this problem using the text retrieval methods that we have learned in this course. As Google Chrome is the most popular web browser, encompassing 65% of browser market share, we plan to create a Google Chrome extension that can index and search through bookmarked pages. We also will allow users to select specific bookmark folders to limit the documents searched as well.

3. **Briefly describe any datasets, algorithms or techniques you plan to use**

We will use the BM25 text retrieval algorithm and evaluate our results on our own bookmarks. We will make our own datasets for evaluation as described below.

4. **How will you demonstrate that your approach will work as expected?**

Following the Cranfield paradigm, we shall create our own test set based off of our own bookmarks to evaluate our results using f-measure, average precision, and other metrics that we learned in lectures. If time permits, we will test our extension on popular text retrieval evaluation datasets.

5. **Which programming language do you plan to use?**

Javascript/HTML/CSS?

6. **Please justify that the workload of your topic is at least 20*N hours, N being the total number of students in your team. You may list the main tasks to be completed, and the estimated time cost for each task**

- Design (2.5 hours) - except idk i guess we do not need it to be too beautifully designed bc then we have to implement it
- Create test cases/evaluation methods (5 hours)
- Research/Identify text retrieval methods (5 hours)
- Frontend - HTML/CSS (15 hours)
- Making datasets (15 hours)
- Backend - (20 hours)
- Create test suite- (5 hours)
- Gather evaluation metrics & Tweak parameters - (10 hours)
- Documentation - (2.5 hours)