

Due at the end of Week 2: Monday, September 26, 6pm. The whole group submits a single pdf on Moodle.

This deliverable is graded out of 10 and contributes 10% to the course grade. Questions 1–3 are worth 8 marks which are attributed for the group: every member will receive the same group mark. Question 4 is worth 2 mark and is attributed individually: every member will receive their own mark.

In this deliverable, all answers are expected to be written in standard English (as opposed to pseudo-code, some programming language, some math formula, etc.). Spelling and grammar mistakes are not penalized as long as the marker can understand what you mean. For each question, I indicate an estimated length of a typical answer. This is not a strict requirement but a recommended guideline.

Question 1 *Overview*

1.1 Propose an internal title for your project. This is not about marketing, advertising, or attracting users. Rather, I want a descriptive title that summarizes what your project is about. It can help you scope the project, and also it can help course staff remember or figure out in 5 seconds what the project is about throughout the term.

Finding out most readable articles in an academic area.

1.2 What is the context and what problem is the project aiming to solve?

The traditional way to learn about a brand-new academic field include reading literature reviews, asking active researchers in this field, and searching for highly cited articles. However, it is difficult to find a high-quality review or to find a suitable expert, while plain citation counts have its shortcomings in measuring academic contributions.

This project aims to develop an adaptive autonomy model that can help people locate what they need to read in a specific research field if they are not familiar with it and want to learn about it quickly. We can also help researchers locate key papers in the developing process within a particular field.

1.3 Describe 2 other projects envisioned and one reason for not choosing to pursue them.

One project is a friend recommendation for social media: Providing a list of users that the current user is potential to interest in and to follow according to their personal profile and existing friend network. We abandoned it because it is a well-studied area and there are lots of matured industrial solutions.

Another project is an optimized voting system that better reflect voters' opinions. We want to develop a voting method that is optimized from plurality voting and ranked voting with combinatorial mathematical methods. We abandoned it because it was difficult to verify the project's actual results since it is probably not a purely computational problem but public affairs/political research.

1.4 Has a variant of this project ever been done before?

Yes.

Some websites visualize the citation networks according to keywords. Some projects use metrics or graph to measure academic influence. However, there's neither KRR-based approach

nor a specific project aim to help researchers to find most readable articles.

[2 marks]

- Give 2 marks if the project seems to make sense overall and the group's answers show careful consideration
- Give 1 mark if the descriptions show sincere effort from the group, but the project is unrealistic or not suited for a KRR approach.
- Give 0 marks if this is empty or it makes very little sense.

Question 2 *Prototyping phase*

The objective of this section is to ensure that the project you are proposing has enough substance that it could serve as for Project B.

2.1 What constitutes an input instance to your problem?

The input instance contains a list of academic literatures, along with their meta data. The meta data include title, authors, tags, reference list.

2.2 How will you get sample data to help test your implementation?

From public academic databases using API such as Crossref and Arxiv API. We can also hand-make some testing data for some special scenarios.

2.3 What constitutes a candidate solution to your problem?

A list of names of Core articles in the specified area or given corpus. Core article is defined by coverage ratio, which measures the influence of an article which is measured by the weighted ratio between influenced works and total articles in the area.

2.4 Presumably, a solution to your problem needs to satisfy several properties or constraints. State two sample constraints that any solution should satisfy.

The returned articles must be relevant to the specified area.
The returned article list should pass a coverage threshold.

[4 marks]

- Each subquestion is worth 1 mark.

Question 3 *Production phase*

3.1 Who are the target users of the system?

Stakeholders includes academic search engine users, researcher to the history of a specific academic discipline, and academic paper database provider.

3.2 Where will the actual data used for the production system come from?

From public authoritative academic databases.
 API for DOI (Crossref).
<https://academia.stackexchange.com/questions/106790/api-to-get-citations-from-doi>
 Arxiv API
<https://arxiv.org/help/api/basics>
 APIs for academic literatures
<https://libguides.mit.edu/c.php?g=1253643p=9182104s-lg-box-wrapper-34205765>

[2 marks]

- Each subquestion is worth 1 marks.

Question 4 *Self Evaluation*

List the contribution(s) of every group member to Project A and self-evaluate every member for up to 2 mark. Example contributions can include but are not limited to brainstorming, determining the problem, its scope, writing up this document and/or other notes.

Answer by filling in the following table. Replace the dummy data with your actual contribution and self-attributed mark (0, 1, or 2).

Name	zID	Contributions	Self-attributed mark
Kunping	z5124475	Brainstorming, writing project definition, reviewing.	2
Shu	z5211077	Brainstorming, information gathering, reviewing.	2
Yiyan	z5183946	Brainstorming, prototyping, reviewing.	2

You are free to set your self-attributed mark however you want, but some might find the following sample quotes helpful.

- “Honestly, I really didn’t do much, but I’ll do better next time (or not).” (0 marks)
- “Sorry, life got in the way and I got too busy with other stuff. I’ve helped a bit and made some difference, but it’s not quite there yet.” (1 mark)
- “I’m satisfied with my contribution. My work might not be perfect or comprehensive, but I did as well as can reasonably be expected from a UNSW student.” (2 marks)

[2 marks]

- No decision is taken by the marker: Each individual gets their own self-attributed mark.