

# Welcome to superproject & Project Proposal

IR Superproject

# In this video

- Enrolment information
- Where to find information about the superproject
- Proposal Draft
- Next meetings in the superproject



Zheng Wang	s4214277	ENGG7813	Qichen Zhang	s4717533	COMP7811
Oliver Carnew	s4432164	REIT4841	Guoqi Li	s4755722	ENGG7813
Abdullah Shahzad	s4480735	REIT4841	Yujie Yuan	s4759342	REIT4841
Krista Bradshaw	s4528514	REIT4841	Bowen Jin	s4763244	ENGG7813
Jiaxiang Li	s4632882	COMP7811	Tianyi Sun	s4763739	ENGG7817
Anqihui Wang	s4637275	COMP7811	Shihuai Wang	s4765194	ENGG7813
Manyang Liu	s4668295	COMP7811	Guan Wang	s4768223	COMP7811
Sirui Wang	s4672966	COMP7811	Yu Zheng	s4774438	ENGG7813
Zian Huang	s4679385	COMP7811	Zirui Tan	s4778753	COMP7881
Pratham Manocha	s4712555	REIT4841			
Shaojiang Li	s4716182	COMP7861			

**8 units:**  
\* **REIT4841**  
\* **COMP7881**  
\* **ENGG7817**  
**6 units:**  
\* **COMP7861**  
**4 units:**  
\* **COMP7811**  
\* **ENGG7813**

# Assessment timeline

COURSE CODES	ENGG4812 (was ENGG4802) METR4912 (was METR4901) REIT4882 (was COMP6804)	REIT4842 ENGG7814 (was ENGG7804) ENGG7818 (was ENGG7808) COMP7812 (was COMP7802) COMP7862 COMP7882 DECO7862
	Assessment	
	First semester of project	
Week 4	<a href="#">Online Academic Integrity Tutorial</a> <i>due by 3pm Thursday of Week 4</i>	
Week 5	<a href="#">Proposal Draft</a> <i>due by 3pm on Thursday of Week 5</i>	
Week 9	<a href="#">Project Proposal</a> <i>due by 3pm on Thursday of Week 9</i>	
Week 11	<a href="#">Seminar and Seminar Participation</a> As scheduled with supervisor <i>due no later than Friday of Week 11</i>	
	Second semester of project (Enrol in the same course code in both semesters)	
Week 11	Not required for undergraduate projects	<a href="#">Conference Paper</a> <i>due by 3pm on Thursday of Week 11</i>
Week 12	<a href="#">Poster and Demonstration</a> As scheduled with supervisor and examiner <i>due no later than Friday of Week 12</i>	
Exam Week 1	<a href="#">Thesis (Electronic)</a> <i>due by 3pm on Monday of Exam Week 1</i>	

8 units:

- \* REIT4841
- \* COMP7881
- \* ENGG7817

6 units:

- \* COMP7861

4 units:

- \* COMP7811
- \* ENGG7813

**This is the only assessment with a difference between # of units!**  
**#4 unit courses - submit a 2 page conference paper**  
**#6 or #8 unit courses - submit a 4 page conference paper**



# So, what do the different # units mean?

- A typical UQ course is 2 units
  - So 4 units project means a course-worth of load per semester
- 8 units per semester is a normal full-time study load
  - So 8 units project means 2 course-worth of load per semester
- UQ says: “It's recommended that you spend approximately 10 hours studying per week for each 2 unit course”
- 4 units project:  $10 * 2 * 13 * 2 = 520$  hours
- 8 units project:  $10 * 4 * 13 * 2 = 1040$  hours

# The IR Superproject

- The GitHub repository for the superproject contains all relevant material: <https://github.com/ielab/IR-Superproject-2023>, and it will be updated on a regular basis
- Common theme: neural rankers
- Different project directions: you are expected to take 1 project direction
- You will need to decide for one project direction for the week 5 assessment, but I am happy for you to change direction (among those identified in GitHub) up to week 7

# Expectation of Progress so far

- You would have watched all videos
- You would have read the generic background material
- You would have read the papers of the research directions and decided on a tentative direction
- You would have identified additional papers related to your direction
- You would have started drafting your proposal
- You have created your own GitHub repository for YOUR project (we will give you instructions soon about adding us to it)

# Proposal Draft

- defines the thesis topic
- presents a review of relevant background material and an assessment of the impact of previous work on the current project
- a brief version of the project plan with a list of milestones, their dates and the required resources needed to complete these
- All background and related material should be appropriately referenced and appear in a bibliography.
- You have been given material in blackboard in weeks 2, 3, 4 that helps you with creating this.



# defines the thesis topic

- This includes both the problem/challenge the work wants to address, and the goals/aims of the project
- Example project direction: reproduce the work — example: <https://arxiv.org/pdf/2112.06400.pdf>
  - The problem challenge is what the original paper aims to address.  
In the example: how to perform effective and efficient relevance feedback with neural rankers
- The goals/aims are:
  - The reproduction of the existing method. In the example: the method by Yu et al. 2021 (ANCE-PRF)
  - The expansion of the original experiments to other datasets/settings/parameters/etc — to show whether the method generalise, or to explore a direction that was not considered in the original work
    - In the example: different hyper-parameter settings (Sec1: “we also aim to further extend that work by considering the factors that affect the training of the ANCE-PRF encoder”), other DRs (Sec1: “generalisability of the strategy underlying ANCE-PRF to other DRs”)

# **presents a review of relevant background material and an assessment of the impact of previous work on the current project**

- Background is: (1) problem/challenge to solve, (2) previous methods related to the problem (e.g. that exhibit the problem, or have attempted to solve the problem)
- In the background videos, we covered a number of methods — some of which you could consider for covering the relevant material for methods.
- Each research direction had one paper provided. In most cases this paper defines the problem
- You SHOULD start from the provided paper and explore all references mentioned in the paper, PLUS you should check who cites the provided paper (use google scholar for this)
- Attempt to identify the key authors in the area — then explore their google scholar profile to check if they have more publication that are relevant
- Of course, you should complement these strategies with actually searching through Google Scholar

**presents a review of relevant background material and an assessment of the impact of previous work on the current project**

- When you talk about previous work, remember it is not just about summarising what they have done
- You want to contrast the previous work with your work
- What does the previous work mean for your work?
- What do you learn from previous work that is important for your work?
- Do the findings in previous work suggest you a direction to follow/not follow?

# Suggested tool for scientific writing: Latex

- A quick way to get to use latex is Overleaf (though I do not like it, it is often quite convenient and quick for starting)
  - I prefer desktop applications, like TexStudio (for Mac)
- Short tutorial on using overleaf Part1: <https://www.youtube.com/watch?v=h4oG33Yfbnc>  
Part2: <https://www.youtube.com/watch?v=uV25oMEuj6A>
- You will save lots of time if you use latex and keep your references in BibTex. Tutorial on using BibTex: <https://www.youtube.com/watch?v=2CoOCC95tHE>
  - Google Scholar has a nice functionality to allow you to export references directly in BibTex; a short tutorial: <https://www.youtube.com/watch?v=EHCIMjxDs-E>



# a brief version of the project plan with a list of milestones, their dates and the required resources needed to complete these

- Two typical resources that you will need are:
- Datasets
  - For directions 1/2/3, check the original papers. You will repeat the experiments on the original datasets, plus potentially you will expand on other datasets. Typical datasets are MS MARCO (inc. TREC DL), ANTIQUE, MQ2007, Robust04. We have these datasets in the research team. Note, for these direction a good dataset to extend the work on will be the DL-typo dataset here <https://dl.acm.org/doi/pdf/10.1145/3477495.3531951>
  - For directions 4/5, Amazon Shopping Queries Dataset; except for NeuCLIR: for this, describe the dataset used in previous year of the task, and check task website for any change.
- Computational resources
  - You should have access to the Rangpur Cluster. We are negotiating with EAIT for some specific instances for you. In the meantime, get describe this resource and get familiar with it.

# Typical structure

- 1. Abstract
- 2. Background and literature review
- 3. Project Plan
- 4. References

# Example proposals

- You can find example proposals at <https://iteethesis-examples.uqcloud.net/search>
- You should consider only those that are related to this area, e.g. Information Retrieval, Machine Learning, Deep Learning. Relevant example proposals
  - On the Efficiency and Accuracy of On-Device Recommendation Systems

Student Number: \_\_\_\_\_ Name: \_\_\_\_\_ Course Code: \_\_\_\_\_

The University of Queensland - School of Information Technology & Electrical Engineering

Mark Sheet for Proposal Draft

Criterion	Pass	Fail
Topic Definition	Clear and concise definition of the research question or problem to be addressed, along with a compelling motivation for why it is important. The topic should be relevant to the field of study and demonstrate originality and potential for contribution.	Unclear or vague research question or problem, with no clear motivation or relevance to the field of study. The topic may be too broad, too narrow, or unoriginal.
Aims	The aims are clearly stated and focused on the overall goals and objectives of the research project and demonstrate a clear understanding of the research problem and its significance.	The aims are vague or unclear, making it difficult to determine the purpose of the research project, or the aims are task-focused rather than outcome-focused, indicating a lack of understanding of the research problem and its significance.
Background	Comprehensive and well-organized review of the relevant background material, providing a good background to understand the topic are presented.	Superficial or incomplete review of the background material, poorly organized, or demonstrates a poor understanding of the topic area
Literature Review	Comprehensive and critical review of the literature related to the research question or problem, demonstrating a strong understanding of the existing research, identifying gaps and limitations.	Incomplete or superficial review of the literature, lacking in critical analysis or coherence, and failing to identify relevant gaps or limitations in the existing research. The review may also be poorly organized.
Selection of references	At least 15 references are presented from reliable sources which have demonstrated relevance to the proposed project.	Some references are unreliable or irrelevant to the proposed project, or there are too few of them.
Milestones	Well thought out milestones, with the associated due dates as well as required resources have been stated. (This section can be brief in the draft version)	Milestones are not well thought out, and not in logical flow, or no associated dates or resources have been identified.
Bibliographic style	The references conform to the IEEE or APA style.	Some references are incorrectly formatted.

Overall:      Pass / Fail [All criteria must pass to pass this piece of assessment]

Comments: \_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

Supervisor Signature: \_\_\_\_\_ Date: \_\_\_\_\_

Supervisor Name: \_\_\_\_\_



# Supervision in the superproject

- You have now all enrolled (the last one today!), so we can start with the group activities
- We would like to have a Q/A session in week 6 to check on your thoughts and doubts after you completed the project proposal draft
- You will have regular group sessions with a tutor
- We want you do NOT communicate with us via email (with the exception of personal issues, e.g. sickness) — but instead that you open an issue on the GitHub repository of the Superproject, so that we can answer there and others can see the answer.