COMP64803:
Introduction to Responsible & Reproducible AI

Mingfei Sun

CDT on Decision-Making for Complex Systems
Department of Computer Science
The University of Manchester

The University of Manchester

In the last lecture

We've learned how to use:

▶ Remote development

▶ Computational Shared Facility (CSF)

▶ Local workstations (4090 RTX)

In this lecture

- ▶ Paper Reproducibility Challenge

- ▶ Project Workflow

- ▶ Expected Outcomes

- ▶ Available Resources

To produce reproducibility reports on very recent papers published between 2022 and 2025 at

▶ Eleven top ML conferences:

NeurIPS, ICML, ICLR, ACL, EMNLP, CVPR, ECCV, AAAI, IJCAI-ECAI, ACM FAccT, SIGIR (list not exhaustive)

▶ Or top ML journals:

JMLR, TACL and TMLR (list not exhaustive)

Paper Reproducibility Challenge (2/3)

Why reproducibility is important?

Essentially, think of your role as an inspector verifying the validity of the experimental results and conclusions of the paper.

► To encourage the publishing and sharing of scientific results that are reliable and reproducible

► To verify the empirical results and claims in the paper by reproducing the computational experiments

► Either via a new implementation or using code/data or other information provided by the authors

Your role will also extend to helping the authors improve the quality of their work and paper.

How our reproducibility challenge work?

You do not need to reproduce all experiments in your selected paper:

▶ For example the authors may experiment with a new method that requires more GPUs than you have access to

▶ You could elect to reproduce only the baseline results

▶ You can check if baseline methods are properly implemented

▶ You can check whether hyper-parameter search is done with the same degree of attention

Paper Reproducibility Challenge (3/3)

How our reproducibility challenge work?

You do not need to reproduce all experiments in your selected paper:

- ▶ For example the authors may experiment with a new method that requires more GPUs than you have access to

- ▶ You could elect to reproduce only the baseline results

- ▶ You can check if baseline methods are properly implemented

- ▶ You can check whether hyper-parameter search is done with the same degree of attention

If available, the authors' code can and should be used

- ▶ The methods described can also be implemented/re-implemented according to the description in the paper

- ▶ You are supposed to detect anomalies in the code, or shed light on aspects of the implementation that affect results.

Project Workflow (1/2)

1. Select a paper and avoid duplicate work
   - Encourage to avoid duplicate reproducibility efforts: search for the papers with open-sourced codes
   - Or papers that have been reproduced by other reproducibility challenges

2. Submit form with the selected papers and codes you find (due by 23 Mar EoD)
   - A form will be live on our Teams channel
   - Title, authors, conference, link to the paper, link to the code
   - Expectations from your supervisors

Project Workflow (2/2)

3. Post reproducibility project on our GitHub wiki pages (due by 20 Apr EoD)
   - ▶ Outline the reproducibility plan, including the experiments you plan to reproduce and milestone dates
   - ▶ Report your initial critics on the paper and the code

4. Submit a report (due by 29 Jun EoD, two weeks after the examination period)
   - ▶ This report should has an abstract (300-400 words) detailing the key findings of your report
   - ▶ Your report should be in ICLR Latex template, within 5-8 pages.
   - ▶ Your report should include assessment comments from your supervisors (as a separate section titled "Supervisor's Assessment")

Timeline

Mar 2026	SEM2W04	23							1
	SEM2W05	24	2	3	4	5	6	7	8
	SEM2W06	25	9	10	11	12	13	14	15
	SEM2W07	26	16	17	18	19	20	21	22
	SEM2W08	27	23	24	25	26	27	28	29
	EVAC 01	28	30	*submit paper selection*					

Apr 2026	EVAC 01	28			1	2	3	4	5
	EVAC 02	29	6	7	8	9	10	11	12
	EVAC 03	30	13	14	15	16	17	18	19
	SEM2W09	31	20	21	22	23	24	25	26
	SEM2W10	32	27	*create a wikipage*					

May 2026	SEM2W10	32					1	2	3
	SEM2W11	33	4	5	6	7	8	9	10
	SEM2W12	34	11	12	13	14	15	16	17
	Ex SEM2W13	35	18	19	20	21	22	23	24
	Ex SEM2W14	36	25	26	27	28	29	30	31

Jun 2026	Ex SEM2W15	37	1	2	3	4	5	6	7
	Ex SEM2W16	38	8	9	10	11	12	13	14
	SVAC 01	39	15	16	17	18	19	20	21
	SVAC 02	40	22	23	24	25	26	27	28
	SVAC 03	41	29	30	*submit your report*				

Expected Outcomes

▶ You should produce a Reproducibility report, describing the target questions, experimental methodology, implementation details, analysis and discussion of findings, conclusions on reproducibility of the paper.

▶ The result of the reproducibility study should NOT be a simple Pass / Fail outcome. The goal should be to identify which parts of the contribution can be reproduced, and at what cost in terms of resources (computation, time, people, development effort, communication with the authors).

▶ You should expect to engage in dialogues with the paper authors.

▶ The top reports are encouraged to be submitted to
`https://reproml.org/`

Available Resources

- ▶ 4090 RTX workstations (will be reserved for this reproducibility challenge for the next 3 months)

- ▶ Computational Shared Facility (CSF)

- ▶ Your supervisors (discuss with your supervisors)

END LECTURE