

MSc in Computer Science (Data Analytics, Artificial Intelligence, Artificial Intelligence - Online)

Capstone Project Guidelines

School of Computer Science, University of Galway

Academic Year 2024-25

Version: 13/01/2025

1 Introduction

The capstone project is a major component of the degree, worth **30** of 90 ECTS. This represents about 600 hours of work. The project is an individual effort; group projects are not possible.

2 Project Main Tasks

- Identify/research a Data Analytics (DA) / Artificial Intelligence (AI) problem
- Identify, describe, and synthesize the state-of-the-art approaches to the problem
- Devise and implement a solution to the problem which has novelty
- Carry out an appropriate (typically experimental) evaluation on the solution
- Write up the problem, state-of-the-art, methodology and implementation, (experimental) evaluation, results and their discussion, and conclusion as a thesis.

The project is in the area of DA/AI. It is a **research project**, not software development (although, typically, the project also involves writing software code). The outcome is a document (the *thesis*) describing your research – not (usually) a viable commercial product.

3 Supervisor Allocation Process

Each supervisor will have a budget of projects, related to their workload (e.g., project supervision on other programmes). Thus, not all students will necessarily be allocated to the supervisor they may prefer. Individual supervisors cannot make commitments to supervise particular students in advance of the final decision by faculty. Allocation of students to supervisors and projects is carried out by faculty to maximize fit and fairness.

4 Supervision

Every student has an academic supervisor, one of the members of University of Galway School of CS faculty. In some cases, there may be a co-supervisor from the same or another School.

In some cases, the academic supervisor may decide that a suitable project should involve an industry partner, in which case there may be a mentor from the sponsoring company in addition to the academic supervisor. A company-mentored project is subject to the same rules as other projects; in particular, the project must be the individual effort of the student, with no group-work involved.

Students may consult <https://www.universityofgalway.ie/science-engineering/school-of-computer-science/people/> and <https://www.universityofgalway.ie/dsi/> for information on supervisors' research interests. The list of academic staff available as possible supervisors will be circulated by the Programme Director.

Students can expect regular meetings with their supervisor in the early stages – once a week or every two weeks are common. The frequency may decrease when the project is “up and running”. Supervisors will typically take annual leave at some point in the project also. In addition to face-to-face meetings, email, phone, online meetings, etc. may be used depending on both parties' preferences.

Supervisors will expect students to be prepared for meetings, including by finding and familiarizing themselves with relevant materials (such as relevant research papers and software tools) and being prepared to summarise them; by working on ideas discussed at previous meetings and being prepared to summarise them; and bringing well-formed questions and own ideas. Students should take notes in meetings. Some students may, e.g., photograph whiteboards after meetings.

Students in the MSc AI Online who are in full-time employment in a related industry may carry out their project in conjunction with their employer if appropriate, subject to University of Galway approval. In these cases, there may be a mentor from the employer. The ideal is to work on a project which is useful to the employer, and thus good for the student's career, as well as fulfilling the requirements of the degree. Students may wish to consider how to demonstrate value to the employer when scoping and writing the project, e.g., choosing certain figures or sections of the text as most appropriate for the employer. There are no extra expectations or workload on the student or the supervisor (relative to full-time).

5 Project Timeline

5.1 MSc DA and full-time MSc AI

- **28th February:** Students submit a 1-page ***Project Definition Document (PDD)*** on Canvas informally describing their proposed project, relevance, possible datasets and availability. A statement of any preliminary reading or other work already underway can also be included. A template for the PDD is provided on Canvas. The PDD needs to be agreed with the project supervisor before submission.

The PDD as such is not graded but a mark penalty applies if no satisfactory PDD is received by the deadline. See Sect. 8 for details.

- **20th June:** Students submit a formal ***Project Proposal (PP)*** document on Canvas and additionally by email to their supervisor. For requirements and further information see Sect. 6.

The PP as such is not graded but a mark penalty applies if no satisfactory PP is received by the deadline. See Sect. 8 for details.

- **21st August:** Students submit their ***Final Thesis*** document on Canvas (via Turnitin) and additionally by email to the supervisor and second marker. A thesis template is provided on Canvas where generic guidelines are given. Individual supervisors should be consulted with respect to any divergence with these guidelines that they might wish to specify. Observe the strict upper limit of 13,000 for the number of words. For further information and submission guidelines see Sections 7 and 10.

- **Early September** (exact dates TBC): Students attend capstone project ***Vivas***. See Section 11 for more information.

Observe that no repeat facility in the same academic year exists for the final thesis.

5.2 MSc AI Online

- **Semester 2 (Year-1), Sunday at end of Week 3:** students provide up to 1-page ***Project Definition Document (PDD)*** to Programme Director informally describing their proposed project, employer relevance, possible datasets and availability, preferred supervisors. Students who choose not to propose their own projects should still provide as much detail as possible on preferred topics/supervisors, and may contact the Programme Director and possible supervisors in advance. Students should then include a very brief summary of any discussion with the preferred supervisor, e.g., including that supervisor's comments on project viability. A statement of any preliminary reading or other work already underway can also be included.

The PDD as such is not graded but a mark penalty applies if no satisfactory Project Definition Document is received by the deadline. See Sect. 8 for details.

- **End of Semester 2 (Year-1):** Literature Review deliverable as part of Research Skills module.
- **30th June (of Year-1):** Students submit a formal ***Project Proposal (PP)*** document on Canvas and additionally by email to their supervisor. For requirements and further information see Sect. 6.

The PP as such is not graded but a mark penalty applies if no satisfactory Project Proposal is received by the deadline. See Sect. 8 for details.

- **31st August (of Year-2):** Students submit their **final thesis** document on Canvas and additionally by email to the supervisor. A thesis template is provided on Canvas where generic guidelines are given. Individual supervisors should be consulted with respect to any divergence with these guidelines that they might wish to specify. Observe the strict upper limit of 13,000 for the number of words. For further information and submission guidelines see Sections 7 and 10.

- There is no viva in the academic year 2023-24 in the MSc AI Online, but this policy might be changed in the next academic year i.e. 2024-25.

Observe that no repeat facility in the same academic year exists for the final thesis.

6 Project Proposal

The formal Project Proposal (PP) needs to describe a research problem to be solved in some detail, including background, motivation, research questions, and proposed research methods.

The PP can be seen as an incomplete draft of the final thesis and should have the same overall chapter structure as the final thesis. It should incorporate a high-quality Literature Review (Background and Related Works), and drafts of the Introduction chapter and the Methodology (or equivalent) chapter. (For students who have taken the Research Skills module, the Literature Review produced earlier may be re-used in the proposal in part or in whole.)

This document should normally be between 11 and 24 pages long (plus references). Supervisors will give feedback verbally, by email or on Canvas.

The template for the final thesis (see Section 7.2) should also be used for the Project Proposal (but normally only a part of the template can already be completed at this time).

The template also contains important guidance for the PP and final thesis' chapters.

Students are strongly encouraged to view the sample theses from previous years provided on Canvas before they start working on the PP and thesis.

Normally, faculty approve the Project Proposal. However, faculty may reject a low-quality proposal and require a complete redesign or a new topic.

7 Final Thesis Document

The document should follow a typical scientific structure (see Section 7.2 / template provided on Canvas). Guideline page counts are given in the template. It needs to clearly distinguish between your primary and secondary research, i.e., between what has previously been accomplished, and what has been accomplished by you.

The number of words in the PDF thesis document (not in the LaTeX source) should be between 8,000 as a minimum and 13,000 as a strict maximum (not including title page, Declaration of

Authorship, Acknowledgements, List of Figures, List of Tables, Table of Contents, Bibliography (References) and Appendixes). **13,000 is a strict upper limit – longer theses may be rejected without review.**

7.1 Writing and Style

The goal of the document is to convince a skeptical, well-informed reader, and to document work done.

The document should be written for other researchers. It should assume the reader knows at least undergrad-level material. It should be written in clear, simple English, with short sentences and little or no “poetic” language. Graphics and tables should be clean, with no frills or decoration.

The best way to develop an appropriate style of writing and formatting a paper is by reading many high-quality research papers. Examples of good-quality theses from previous years are also available on Canvas.

A draft which has not been edited will rarely achieve quality. “The best writing is rewriting.” - E B White. “My pencils outlast their erasers.” - Vladimir Nabokov. Supervisors will give feedback on writing. The best way to receive useful feedback is to edit your own drafts first, then request feedback on specific sections and/or on overall structure – the earlier, the better.

The document should be focused on relevant material, with no padding and little repetition. (But it is appropriate to have some repetition between the Abstract, Introduction, and Conclusions; and some repetition where main points in the Introduction are expanded elsewhere.)

7.2 Thesis Template

We provide a thesis template in LaTeX format. The .zip is downloadable from Canvas. It contains instructions for use. LaTeX is recommended, but students who prefer other software are free to use it. They need to mimic the thesis template format. The Project Proposal and final thesis should both adhere to this format and should both be submitted in PDF format. So, the proposal can be seen as a first (incomplete) draft of the thesis. Also see Sect. 10 (Thesis Submission).

The LaTeX template also contains important guidance for the PP and final thesis chapters.

8 Grading

The thesis will be evaluated by two internal markers (normally the supervisor and another member of staff from the School of Computer Science or an affiliated institute). It will then be discussed at a meeting of all markers. Grading is overseen by an external examiner from outside the University of Galway.

For each milestone document (Project Definition Document (PDD), Project Proposal (PP)) which is not received on Canvas by the respective deadline (see Sect. 5), the project module % mark will be reduced by 5. These penalties also apply if the PDD or PP submitted by the deadline are not satisfactory.

Example: If the thesis mark is 62%, but both the PDD and the PP were submitted late, the overall mark of the capstone project module will be $62 - 5 - 5 = 52\%$.

However, no mark penalties apply if the late submission of the PDD or PP occurred for certified reasons outside the control of the student, such as illness (with a supporting medical certificate or counsellor's letter).

9 Plagiarism

All submitted work (including the final thesis, PDD, PP, data sets, and software code) needs to be entirely the student's own work and written in the student's own words, except where explicitly stated otherwise in the respective document using the academic discipline's accepted norms and citation style.

Each thesis will be automatically screened to detect plagiarism, but this is just one of the methods we use. Students who try to "trick" automated plagiarism detectors are guilty of attempting to subvert the University's examination procedure. This can be detected by other methods.

Observe that low "similarity scores" obtained from plagiarism detection tools do not mean that a thesis does not contain plagiarism. It remains the full responsibility of the student to ensure that

all materials they used are properly attributed to the exact sources.

9.1 Accidental Plagiarism

Accidental plagiarism means plagiarizing content unintentionally. Accidental plagiarism is plagiarism and is subject to the same procedures and penalties as other forms of plagiarism.

9.2 Examples of Plagiarism

- Use of material created or provided by another person or agency (such as a “ghostwriter” or “essay mill”) or by an Artificial Intelligence;
- Copying the work of another student or individual, with or without that person’s consent;
- Submission of a student’s own work for credit in more than one course;
- Copying from a website, book, journal, conference proceedings, workshop notes, tutorial, thesis, report, preprint, manual or other document, without proper citation;
- Use of material from a website, book, journal, conference proceedings, workshop notes, tutorial, thesis, report, preprint, manual or other document, without an appropriate reference to the exact source;
- Use of ideas, words, concepts, methods, approaches, algorithms, data, code, images, figures, software, APIs, web content or other materials created by someone else without an appropriate reference to the exact source;
- Use of a direct quotation, not included in quotation marks and properly attributed to the original source;
- Paraphrasing of the work of another person without properly attributing to the original source;
- Close paraphrasing of the work of another person, such as, for example, where only superficial changes are made to the original wording or sentence structure;

9.3 Academic Integrity Policy

The University’s Academic Integrity Policy and appendix are available at the following link (under “Learning, Teaching and Assessment”):

<https://www.universityofgalway.ie/registrar/policies-forms/#tab2>

9.4 Declaration of Authorship

Your final thesis must contain the following Declaration of Authorship (on a page of its own directly after the title page). The Declaration of Authorship must be signed by you.

I, <your name>, hereby declare that this thesis, titled “<thesis title>”, and the work presented in it are entirely my own except where explicitly stated otherwise in the text, and that this work has not been previously submitted, in part or whole, to any university or institution for any degree, diploma, or other qualification.

<Your signature>

10 Thesis Submission

Upload your final thesis in PDF format on Canvas (via Turnitin, a plagiarism detection service) by the deadline (see Section 5), and **additionally** send it, directly after submission on Canvas, **by email to your supervisor and to your second marker.**

The name of your second marker will be announced by the Programme Director ahead of the thesis submission deadline.

Also, students need to send by email a *draft* in PDF format to (only) the supervisor **one week in advance of the thesis submission deadline.**

The submission must also include a link to a Web location (e.g., a private or public GitHub repository) where the data and code can be found, except where agreed otherwise in advance with the supervisor, e.g., because of sensitive data.

Remember that the final thesis needs to include a signed **Declaration of Authorship** as specified in Section 9.4.

11 Viva

Every student in the full-time MSc AI and MSc Data Analytics is required to attend a short viva voce (“viva”) with the two project markers (normally the project supervisor and a second marker

from the School of CS or an affiliated institute, see Section 8) after submission of the thesis.

Not attending the viva, without a valid justification, will result in reduced marks and may imply failure of the capstone project.

The viva dates and times will be announced by the Programme Director.

The purpose of the viva is to give the markers the opportunity to assure themselves that the work presented in the thesis is the student's own, to improve the markers' understanding of the contribution the student made with their project, and to give the student the opportunity to clarify any matters raised by the markers.

The duration of the viva is normally about 20-25min. The focus of the viva is on questions about the project and about the student's contribution. A formal presentation is not part of the viva, but the student is expected to provide a brief overview of their project at the beginning of the viva.

The viva takes place at a date and time set by the Programme Director, unless the two markers and the student agree on a different date/time. If the student is unable to attend at the determined date and time for an attested reason beyond their control (such as illness attested by a medical certificate or counsellor's letter), the student is required to inform both markers as soon as feasible before the viva and to request a new viva date and time from them.

The student must contact both markers at least two weeks before the viva date, to confirm date, time and venue of the viva. The student is also required to email the final version of the thesis to both markers directly after submission on Canvas (see Section 10).

The viva normally takes place in-person on campus (in the office of the project supervisor, unless advised otherwise by the supervisor). However, if the student or a marker is not available for an in-person viva for an important reason (for example, if the student does not live in Ireland anymore at the time of the viva), the viva takes place online.

In case the viva is online, the student needs to send, at least two weeks before the viva date, a Teams calendar invite for the viva (including the Teams meeting link) to **both** markers (not only

to the supervisor).

If the viva takes place online, it is the responsibility of the student to ensure that the student has the necessary technical facilities (such as a computer with camera and microphone, a reliable broadband connection, and access to Teams) available.

12 Publication

For faculty, publication is a core goal of research. A thesis which we evaluate as publishable receives a high grade. Students can optionally demonstrate publishability by extracting a submission-ready research paper from their thesis. This would typically be 8-15 pages in a much more compressed and focused style, targeted at a specific journal or conference. Students should discuss the possibility with their supervisor before engaging in this work.

For some projects, it may be appropriate to provide Latex source, code and data to the supervisor to allow for further experimentation and writing either by the supervisor or by a future student – standing on the shoulders of giants.

13 Intellectual Property

The copyright for documents, code and results produced by the student remains with the student. (Students in employment may need to refer to their employer and/or contract for clarification on copyright.)

Students should discuss IP issues with their supervisor if the supervisor is contributing ideas, data, or any other IP to the project.

Our strong preference is for projects involving open, anonymised or synthetic data such that the student is entitled to publish their work in the thesis and in scientific venues. Note that such publication does not necessarily require the data to be made public.

Where company data is used, workarounds may be needed to preserve sensitive data. In some cases, anonymised or synthesized data can be used to avoid exposing students or supervisors to sensitive data. If a sponsoring company wishes, they can request that the students not include any sensitive information in the project outputs, so long as this doesn't prevent the students from describing their findings. For example, the students could describe their findings in detail while

omitting mention of specific customer names or monetary amounts. If appropriate, the supervisor, the students and the employer would agree broadly at the outset what types of data can be included, and agree a timetable by which the students will provide representative drafts of documents or slides, and dates by which the partner will respond with confirmation that they can be published.

University of Galway faculty cannot in general sign Non-Disclosure Agreements for MSc projects. Normally, thesis documents are considered as published material, and may be (e.g.) copied to other faculty and future students, or deposited in the Library, or on a web page. Students (especially those in employment, carrying out projects in conjunction with their employer) may request an embargo. This means that we undertake not to deposit the document (and associated code, data, presentation slides, etc.) in the Library, or on our websites, or to allow their contents to become known outside the School of CS faculty.

14 Research which requires approval/consent

Certain types of research, for example Health Research or research which involves personal data of individuals, require consent of participants and/or approval of certain committees and/or may be subject to other requirements, to comply with regulations such as GDPR. Your project supervisor will be able to advise.

University of Galway Data Protection Website:

<https://www.universityofgalway.ie/data-protection/>