2023

Bachelor's Thesis Seminar In Computer Sciences And Information Technology

Matti HAAVISTO
Tampere University, Finland, matti.haavisto@tuni.fi

Hannu-Matti JÄRVINEN
Tampere University, Finland, hannu-matti.jarvinen@tuni.fi

Terhi KILAMO
Tampere University, Finland, terhi.kilamo@tuni.fi

See next page for additional authors

Recommended Citation

Haavisto, M., Järvinen, H.-M., Kilamo, T., Ovaska, S., & Poranen, T. (2023). Bachelor's Thesis Seminar In Computer Sciences And Information Technology. European Society for Engineering Education (SEFI). DOI: 10.21427/5F0D-8J34

BACHELOR'S THESIS SEMINAR IN COMPUTER SCIENCES AND INFORMATION TECHNOLOGY

M. Haavisto
Tampere University, Tampere, Finland
ORCID 0009-0005-8915-9708
H.-M. Järvinen
Tampere University, Tampere, Finland
ORCID 0000-0003-0047-2051
T. Kilamo
Tampere University, Tampere, Finland
ORCID 0000-0002-9561-1116
S. Ovaska
Tampere University, Tampere, Finland
ORCID 0000-0002-7318-4766
T. Poranen*
Tampere University, Tampere, Finland
ORCID 0000-0002-4638-0243

Conference Key Areas: Curriculum Development, Fostering Engineering Education Research

Keywords: Bachelor's thesis, seminar work, thesis supervision

ABSTRACT

A Bachelor's thesis is typically an individually written literature review on a scientifically relevant topic. Additionally, some theses also describe empirical work or report an experiment. Firstly, we introduce how Bachelor's theses are supervised in a joint thesis seminar for Computer Sciences and Information Technology at our university. The thesis seminar is organized three times a year. It consists of six small group meetings led by a supervisor and contains compulsory pre- and post-assignments and active peer discussions. In 2022, there were in total of 187 students participating in the spring, summer and autumn seminars. Secondly, we give an overview of the 98 completed theses. We classify the theses using ACM's Computing Classification System and analyze keywords, the number of references and some other bibliometrics to learn about the students and the potential effects of their different study orientations. We also analyze 14 theses that reported practical work, like the implementation of an algorithm or using existing software tools. The main result of our work is to give a research-based view on the supervision of Bachelor's theses, the organisation of the thesis seminar, and the bibliometrics of the completed thesis.

*Corresponding author

1 INTRODUCTION

In 2019 the two universities in Tampere (University of Tampere, and Tampere University of Technology) were merged into Tampere University (Tampere University 2018). The new multi-disciplinary university is the second largest in Finland with almost 20,000 students in bachelor's degree and master's degree study programs. The universities had degrees in Computer Science (CS) and Information Technology (IT) with good records of graduates throughout the years. While the separation between the academic fields is profound, and the graduates will still get their diplomas in BSc or BSc in Technology, it is most visible in other than the major field of study, since the degree programmes share most of the software-related studies. Since the merger, teaching has also been harmonised between different cultural and skill backgrounds to cope with the ever-increasing numbers of students, the demands of the pandemic and the pedagogical changes that all these have brought.

In this paper, we focus on Bachelor's theses and the thesis seminar where they are supervised. The thesis is one of the final tasks for the student before graduation, usually in the third or fourth year of study. It would be in the interests of the faculty to get more graduates yearly since part of the funding is based on the number of Bachelor's and Master's degree diplomas awarded, and the funding model emphasizes on-time graduates - those who complete their degree in three academic years.

The aim of our research is to look into the joint thesis seminar to see how well it serves the degree and learning goals of the CS and IT degree programmes. Our research question is:

1. What kind of theses are completed, and do they indicate differences between CS and IT students?

Our analysis is based on the accepted theses available in the public thesis database of the university as well as the enrollment statistics for the 2022 calendar year.

The rest of the paper is structured as follows. In Section 2 we describe the background of the Bachelor's thesis in the computing curricula. In Section 3 we focus on our implementation, its learning outcomes and seminar structure. Section 4 shows the results of our analysis of seminars in 2022. The results are discussed in Section 5 which concludes the paper.

2 Bachelor's thesis in computing curricula

A Bachelor's degree includes either a thesis or a large-scale final project (e.g., a capstone project). In Europe, a thesis seems to be common at universities, and final projects occur more often at universities of applied sciences but this is not a rule. The length of a European Bachelor's degree is three or four years as stated in the Bologna Declaration (European Higher Education Area 1999). In the USA, a capstone project is common instead of a thesis seminar (Blumenthal 2022), especially in the context of industry projects (ACM 2020). The U.S. Bachelor's programmes are mostly four years.

Bachelor's theses are also closely tied to the teaching of research methods, as a properly written thesis contains a literature review that ties the work into existing knowledge in the research area. Koppelman et al. (Koppelman, Dijk, and Hoeven 2011) reported in their case study findings from an undergraduate research course. One of their results was that students feel better prepared to conduct research in graduate programs after the research course. Holz et al. (Holz et al. 2006) point out the large variety of research methods in computing and discuss how they should be taught to students. In our study programs research methods are discussed in the MSc level. However, the literature review approach naturally involves getting familiar with the research methods applied in the literature of one's study field.

2.1 Bachelor's theses in higher education programmes

In Finnish BSc degree programmes, a thesis of a minimum of six and a maximum of ten credits is required by the legislation (Finlex 794/2004 2004). The credits follow the European credit system ECTS (European Commission 2015) where one credit is $26\frac{2}{3}$ hours of student work. Together the seminar and the thesis bring 10 ECTS. The full academic year is 60 credits, and the Bachelor's degree 180 credits in the relevant areas of this paper.

2.2 Bachelor's thesis guidelines

The Bologna process does not specify anything about the Bachelor thesis. Likewise, a thesis is not mentioned in the ACM Computing Curricula 2020 (ACM 2020) though it encompasses BSc programmes in Computer Engineering, Computer Science, Cybersecurity, Information Systems, Information Technology, and Software Engineering. Some earlier computing curricula recommendations by the ACM require a final project, but only the Chinese version of the information technology curriculum includes a graduation thesis (ACM 2017). In computer engineering, one example curriculum mentions a final individual project that includes a thesis (ACM 2016), and in computer science, one example curriculum recommends project courses with a note that a "Reading, Research, or Thesis course" is not enough (ACM 2013).

The only example curriculum containing a Bachelor thesis is the Chinese four-year version of the information technology curricula. It describes the contents of the thesis as follows: "Students do literature translation, literature survey, opening report, system design and development, thesis writing and defending; students acquire scientific research ability, system design and development ability, develop a basis for future work." (ACM 2017) In this respect, the ACM reflects the practices in the USA. However, even if the theses are common in Europe, the Bologna process does not mention a Bachelor-level thesis.

Universities have published local guidelines for bachelor theses, which may be subject-specific or aimed at the whole university. The study guide of Tampere University falls into the latter category. The objectives of the thesis are described as follows.

> With a bachelor's thesis, students demonstrate their ability to apply their acquired knowledge and skills, engage in scientific or artistic thinking and activities and communicate effectively, both orally and in writing, in their mother tongue. Students typically attend a bachelor's thesis seminar while working on their bachelor's thesis. (Tampere University 2019)

3 BSc thesis seminar in computer sciences and information technology

Our paper discusses theses written as a part of a three-year Bachelor's programme at Tampere University. The BSc theses in Computer Sciences and Information Technology are completed by taking a semester-long thesis seminar course. The seminar is arranged three times a year - Autumn, Spring and Summer.

3.1 Learning outcomes

In the seminar, there are two main learning outcomes that are common to students of both degree programs.

1. Learn how to do a small research work with the structure of a common research paper.
2. Learn to write scientific text.

However, the learning outcomes for the Bachelor's thesis in Computer Sciences and in Information Technology are verbalized separately for each in the current curriculum. This is due to the histories of two separate universities, and will undergo further unification for the

next curriculum period 2024-2027. Still, when coded together, the core learning outcomes are common for the two degrees both aiming for a completed BSc thesis required in the degree. Together, they include the following. After completing their thesis:

- The student has practised writing a thesis. They know how to design and write a thesis and how to take the academic audience into account.
- The student has experience in searching for and reading scientific and technical papers and writing their results comprehensively. They are able to systematically search for information and recognize the sources relevant to their field. The student knows how to evaluate and utilize sources of information in their thesis and is able to exclude sources not relevant to their work.
- The student has practised interaction with other professionals and knows how to give and receive scientific critique. Their comments are useful and constructive. The student is able to evaluate comments and handle them appropriately.
- The students know how to analyse the key elements of a research problem and their relationships. They understand the nature of the scientific writing process and can apply it in practice.

In the learning outcomes, differences can be identified in the target audience of the thesis. IT highlights the technical engineering audience while CS targets a more holistic professional audience. CS emphasizes the ability to form and defend independent views regarding a research problem while IT takes a more practice-driven approach of verbalizing the learning outcome through the ability to present the student's work and to act as an opponent. CS mentions the ability to address ethical concerns explicitly. IT addresses the ability to take the reader's needs into account.

Between the two degree programs, the main differences in requirements can still be associated with academic writing studies. Whereas students in Information Technology complete an "Academic Writing" module integrated into the seminar, the students in Computer Science have a separate scientific writing course with a scope of 2 ECTS. The reason for such imbalance lies in the differences of the study fields. As the engineering students have natural sciences (mathematics, physics, chemistry) as required studies, their study program is tightly packed into 180 ECTS to match three years of study, and for them, the integration of writing counselling to the seminar itself does not produce credits.

As a form of thesis, a literature review is recommended but constructive research is allowed. In all cases, a thesis is based on a literature review on the topic. Planning and carrying out empirical or constructive research often takes long, thus causing more student workload than expected for the seminar, and possibly lengthening the process of graduation. All students enrol in empirical group projects in their studies to gain practical experience, and in our curricula, these projects are separate from the thesis seminar.

The minimum length of the thesis is 10 pages and it should not exceed 25 pages. It can be written in Finnish or English.

3.2 Thesis evaluation

The theses are evaluated with a five-level grading scale: 1 (sufficient), 2 (satisfactory), 3 (good), 4 (very good) and 5 (excellent). In principle, the thesis might also receive a failed grade, but the supervision process prevents this as the supervisor does not allow the student to make an official submission before the thesis reaches an acceptable level. After the student has submitted the thesis for assessment, the supervisor has 21 days to evaluate the thesis by writing a statement. The statement is then approved by the responsible teacher before it is

forwarded to the administration.

The evaluation is based on nine criteria: i) Topic, objectives and thesis title; incl. research question. ii) Structure; structuring the topic and logic of the structure. iii) References; quality, quantity and usefulness, citation practices. iv) Conclusions, achievement of goals and criticality. v) Language, text fluency and appearance (incl. figures and tables). vi) Self-initiative and consideration of feedback. vii) Seminar work activity. viii) Presenting the work and being an opponent in the final seminar meeting. iv) Completing the thesis on schedule. The final grade is not necessarily the average of the criteria, the emphasis depends on the topic, content, and other relevant factors.

3.3 Seminar structure and supervision

Arranging the seminars requires lots of coordination work by a team of responsible teachers who handle general arrangements before, during and after the seminar. Each seminar instance needs to have both students that enrol with initial ideas of their interests and a preliminary topic for the thesis, and supervisors that are experts in the topics chosen by the students. In addition to the content experts, the students also meet with information search experts of our University Library, and Academic Language teachers.

At the start of the seminar, the enrolled students are divided into small groups of 4 to 8 members based on their initial topics. Every small group is led by an experienced teacher or a professor. The groups follow the meeting agendas described in Table 1. The meetings are held two, three or four weeks apart, depending on the phase of the seminar. We have Moodle (Moodle LMS 2023) as our learning management system, hosting both timed discussion forums for the small groups as well as links to shared lecture videos and practical advice on the required tasks.

Table 1: Seminar structure and group meetings

Before the seminar	Enroll, indicate individual interest areas and initial topic.
Group meeting 1	Meet the supervisor and group members. General ideation.
Group meeting 2	Make first mind map, write motivational paragraphs on the topic.
Group meeting 3	Information search with keywords to find articles, make a concept map.
Group meeting 4	Combine the parts into a thesis outline, indicate missing parts.
Group meeting 5	Write missing parts into an almost complete thesis document.
Group meeting 6	Prepare a presentation. Give detailed feedback as an opponent.
After the seminar	Finalize the thesis and submit it.

The full duration of the seminar is roughly four months, and the students are expected to finish their theses within one to six months after the end of the seminar. All theses are checked against plagiarism and stored in the document repository of our university (Tampere University 2023).

The seminar is started by more students than finish it (Table 2). Similarly, the number of completed theses is lower than the number of completed seminars. Partly the low completion rate is due to the administrative delay of several weeks between the submission of the thesis and it becoming available in the document repository. This is why we extended our data collection until the end of January, as indicated in (Table 2).

4 BSc theses in computer sciences and information technology

Our analysis is based on the 98 accepted theses acquired from Trepo (Tampere University 2023). 90 theses were written in Finnish, and 8 theses in English. Most of the theses, 93,

Table 2: Seminars, participants, and completed theses 1.1.2022-31.1.2023.

Start time	Enrolled	Completed seminar	Completed thesis
Spring 2022 seminar	89	65	57
Summer 2022 seminar	32	21	14
Autumn 2022 seminar	66	44	27
Total	187	130	98

are publicly available online, and 5 theses have only abstracts and keywords online. These five theses can be read at the thesis point in the university library.

4.1 Keyword and topic analysis

We categorized the thesis topics using ACM's Computing Classification System (ACM 2012). The ACM CCS is first briefly explained to the students as they start their information search for thesis topics. We have found it a good starting point in search for up-to-date references. The categorization of our thesis sample (98 theses) is shown in Table 3.

Table 3: Theses categorization using ACM's Computing Classification System.

CCS category	IT students	CS students	All
Applied computing	3 (5.2%)	4 (10.0%)	7 (7.1%)
Computing methodologies	4 (6.9%)	5 (12.5%)	9 (9.2%)
Computer systems organisation	5 (8.6%)	0 (0.0%)	5 (5.1%)
Hardware	2 (3.4%)	0 (0.0%)	2 (2.0%)
Human-centered computing	5 (8.6%)	12 (30.0%)	17 (17.3%)
Information systems	10 (17.2%)	2 (5.0%)	12 (12.2%)
Mathematics of computing	0 (0.0%)	0 (0.0%)	0 (0.0%)
Networks	0 (0.0%)	0 (0.0%)	0 (0.0%)
Security and privacy	6 (10.3%)	5 (12.5%)	11 (11.2%)
Social and professional topics	4 (6.9%)	3 (7.5%)	7 (7.1%)
Software and its engineering	14 (24.1%)	7 (17.5%)	21 (21.4%)
Theory of computation	5 (8.6%)	2 (5.0%)	7 (7.1%)
Total	58 (59.2%)	40 (40.8%)	98 (100%)

Mathematics of computing students complete their thesis in a separate seminar with students of mathematics and statistical data analytics. Similarly, students of Networks participate in the BSc thesis seminar in Electrical Engineering (EE). This is due to Networks as a major being common to IT and EE, and as the students are grouped based on their initial topics, it makes sense to keep the Networks students together. Hence these two categories do not have students in our analysis.

The distribution of IT and CS students' thesis topics is in line with the profiles of the degrees with IT being geared toward software engineering, information and computer systems while CS has more of an emphasis on human-technology interaction. Societal aspects and information security are equally present in both. Each student chose up to six keywords to describe their thesis work, and we combined those keywords into loose semantic clusters that are presented in (Table 4) according to their size.

Specific development techniques and tools (such as React, Django, Java, Javascript, C++, Python, ...) often appeared in the keyword lists but were left out of this analysis.

Table 4: Clusters of most common keywords in the theses.

	IT students	CS students	All
1. software engineering, agile, projects	15	7	22
2. machine learning, neural networks	9	6	15
3. algorithms, computation	6	9	15
4. information security, cybercrime	3	11	14
5. programming, web, mobile	4	7	11
6. accessibility	2	9	11
7. data science, data bases	5	5	10
8. usability, user experience	3	6	9
9. social media	3	3	6
10. recommender systems	3	3	6

4.2 Thesis bibliometrics

Thesis bibliometrics for a number of pages, words and references are listed in Table 5. The minimum page length required is 10 and the minimum word count for a thesis to be acceptable is 3000 words. The thesis should have a minimum of 10 references with 15 as the recommended average. The average number of pages in the theses was 18.48. The shortest thesis was 11

Table 5: Average number of pages, words and references of CS and IT theses.

	Pages (min/ave/max)	Words (min/ave/max)	Ref. (min/ave/max)
CS students	11 / 17.60 / 24	3215 / 4986.74 / 7920	12 / 19.45 / 33
IT students	12 / 19.09 / 30	3105 / 5099.04 / 8303	10 / 18.97 / 44
All	11 / 18.48 / 30	3105 / 5053.15 / 8303	10 / 19.16 / 44

pages and the longest was 30 pages. An average CS thesis was 17.6 pages and an average IT thesis was 19.09 pages long. IT students seem to write longer theses as the shortest CS thesis was 11 pages and IT 12 pages, and the longest CS thesis was 24 pages while the longest IT thesis was 30 pages. However, the thesis template in IT renders more pages and thus the number of words gives a better comparison point.

The shortest thesis was 3105 words (IT) and the longest, also in IT, 8303 words. The longest CS thesis was 7920 words. On average the theses were 5053.15 words long. The IT average, 5099.04, was a bit higher than the CS average of 4986.74. However, the differences are not significant.

The average number of references was 19.16 (19.45 in CS, 18.97 in IT). The smallest number of references was in an IT thesis, 10. In CS, the lowest amount of references used was 12. Similarly, the largest number of references was found in an IT thesis, 44 with 33 references respectively in CS. The students are expected to discuss the scientific quality of the studies they write about in the thesis. Our reference counts do not take into account the quality or the publication forums of the references. That kind of analysis might indicate more differences between the student cohorts, the IT students being more oriented towards practice.

4.3 Theses with practical experiments

All the analyzed theses contain a literature review on the topic. There were 14 theses that reported a practical study. These theses can be classified into three main categories (the number of theses in the category is given in the parenthesis):

A) Implementing an algorithm or designing a method or model, or designing a challenging and useful example. (7)

B) Applying existing tools to evaluate something. (3)

C) Collecting data with different methods from different sources and then comparing the collected data. (4)

The seven theses in category A described often an implementation of a commonly known algorithm, like the A* search algorithm or a method to generate random numbers. In most cases, they also contained a small-scale evaluation. Some students published their source codes in a public code repository.

Theses in category B (3 theses) reported applying some existing software tools to experiment, collect and compare data. For example, a student studied accessibility reports generated with WAVE on some web e-commerce sites.

Then there were 4 theses (category C) that reported a comparison based on data that was collected from the literature or from other sources, like websites. For example, in one thesis data was collected using a questionnaire form, and in another thesis, data was collected from GitHub, Stack Overflow and LinkedIn.

5 DISCUSSION AND CONCLUSION

To answer our research question, we did not find significant differences between the CS and IT students' theses.

Minimum requirements have an impact on the overall quality of the thesis. Mostly the students wrote theses exceeding the minimum requirements and the average thesis was along the lines of the recommended, not the minimum. While there are differences in the topics and research methods, having common quality guidelines can be viewed as beneficial.

We sought for causes behind the large number of dropouts (see Table 2). A closer look into the dropout statistics reveals some facts. In 2022, altogether 7 students enrolled twice or even three times in the seminar. From them, only one student has finished the thesis after the second seminar by the time of writing. One reason for this may be that many students in the Computer Science and Computing fields are more qualified in writing Python, C++ or similar code segments than in natural languages. For them, writing the thesis seems difficult and beyond the skills of the student. Similar results pointing out the importance of beliefs of self-efficacy are brought up e.g. in (Blankenstein et al. 2019).

A second cause for discontinuing the seminar is beyond the scope of the seminar itself. Despite our efforts on the seminar arrangements, the thesis is often postponed as far as possible. After graduation most students continue in their Master's degree program and there is no clear cutoff point between the programs. To complicate the matter, Finnish students do not pay study fees in higher education, and they do not see graduation as high on their priority list. Many of them carry out outside jobs to support living or work in an internship to gain experience in the study field, leaving less time to focus on their studies. The national BSc students' survey findings (Education Statistics Finland 2023) for 2022 found that the top three main causes of delays in studies were the lack of motivation to study, health-related reasons and working. The lack of motivation was the leading cause and was on the rise from the previous year. While not directly in the hands of the thesis seminar alone, these aspects are important to take into account and address when planning the seminar.

To wrap up, finding a suitable topic and the motivation to spend time on it to write a thesis requires resilience from the students. A comforting fact is that while thesis writing may take time over the intended time frame, most students eventually get their thesis completed, and the seminars enable effective focusing on thesis writing.

References

ACM. 2012. *ACM Computing Classification system.* Visited 23.2.2023. https://dl.acm.org/ccs.

———. 2013. *ACM Computer Science Curricula 2013.* Visited 7.3.2023. https://www.acm.org/binaries/content/assets/education/cs2013_web_final.pdf.

———. 2016. *ACM Computer Engineering Curricula 2016.* Visited 7.3.2023. https://www.acm.org/binaries/content/assets/education/ce2016-final-report.pdf.

———. 2017. *ACM Information Technology Curricula 2017.* Visited 7.3.2023. https://www.acm.org/binaries/content/assets/education/curricula-recommendations/it2017.pdf.

———. 2020. *ACM Computing Curricula 2020.* Visited 7.2.2023. https://www.acm.org/binaries/content/assets/education/curricula-recommendations/cc2020.pdf.

Blankenstein, F. M. van, N. Saab, R. M. van der Rijst, M. S. Danel, A. S. B.-v. den Berg, and P. W. van den Broek. 2019. "How do self-efficacy beliefs for academic writing and collaboration and intrinsic motivation for academic writing and research develop during an undergraduate research project?" *Educational Studies* 45 (2): 209–225. https://doi.org/10.1080/03055698.2018.1446326.

Blumenthal, R. 2022. "Alignment among Normative, Prescriptive, and Descriptive Models of Computer Science Curriculum: The Effect of ABET Accreditation on CS Education." *ACM Trans. Comput. Educ.* (New York, NY, USA) 22, no. 3 (July). https://doi.org/10.1145/3513141.

Education Statistics Finland. 2023. *Finnish Bachelor's Graduate Survey.* Visited 4.4.2023. The URL is shortened. https://urly.fi/389p.

European Commission. 2015. *European Credit Transfer and Accumulation System.* Visited 3.3.2023. https://education.ec.europa.eu/education-levels/higher-education/inclusive-and-connected-higher-education/european-credit-transfer-and-accumulation-system.

European Higher Education Area. 1999. *The Bologna Declaration of 19 June 1999.* Visited 7.3.2023. http://www.ehea.info/page-ministerial-conference-bologna-1999.

Finlex 794/2004. 2004. *Government Decree on University Degrees, Section 9 (2013/1039), Structure of a Bachelor's degree.* Visited 21.2.2023. https://www.finlex.fi/en/laki/kaannokset/2004/en20040794.

Holz, H. J., A. Applin, B. Haberman, D. Joyce, H. Purchase, and C. Reed. 2006. "Research Methods in Computing: What are they, and how should we teach them?" In *Working group reports on ITiCSE on Innovation and technology in computer science education,* 96–114. New York, NY, United States: Association for Computing Machinery.

Koppelman, H., B. van Dijk, and G. van der Hoeven. 2011. "Undergraduate Research: A Case Study." In *Proceedings of the 16th Annual Joint Conference on Innovation and Technology in Computer Science Education,* 288–292. ITiCSE '11. Darmstadt, Germany: Association for Computing Machinery. https://doi.org/10.1145/1999747.1999828.

Moodle LMS. 2023. *Moodle learning platform.* Visited 4.4.2023. https://moodle.org/.

Tampere University. 2018. *Together we are greater.* News 13.12.2018. https://www.tuni.fi/en/news/together-we-are-greater.

Tampere University. 2019. *Bachelor's thesis*. Visited 21.2.2023. https://www.tuni.fi/en/students-guide/handbook/uni/studying-0/thesis/bachelors-thesis.

———. 2023. *Trepo, Open institutional repository*. Visited 3.3.2023. https://trepo.tuni.fi/.