Research Seminar on Software Engineering (3 ECTS)

Dietmar Pfahl
dietmar.pfahl@ut.ee
Seminar Goal

• To help you deliver a high-quality master thesis on time
  • Target: Defence in June 2023

• Focus on **Goals** and **Methodology** (to achieve the goals)
• In the Spring seminar you already learned how to do a (superficial) literature review OR a mapping study
Goal of today’s session

At the end of this session you should ...

• Understand the structure and form of your MSc thesis report
• Know about types of theses
• Know about types of goals/research questions
• Understand the schedule of the seminar
• Understand the deliverables that you have to submit
• Understand the criteria for successfully completing the seminar (‘pass’ grade)
Thesis Template


• The typical structure of a thesis consists of the following components:
  • Title page
  • Information sheet
  • Table of Contents
  • Introduction
  • Terms and notions (optional)
  • Chapters
  • Summary
  • References
  • Appendices (if relevant)
  • License
Thesis Template

• Link: https://cs.ut.ee/en/content/thesis-deadlines-and-guidelines

• The typical structure of a thesis consists of the following components:
  • Title page
  • Information sheet
  • Table of Contents
  • Introduction
  • Terms and notions (optional)
  • Chapters
  • Summary
  • References
  • Appendices (if relevant)
  • License

<table>
<thead>
<tr>
<th>Section</th>
<th>Why do it? (Context / Motivation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>What to achieve? (Goals)</td>
</tr>
<tr>
<td>Background</td>
<td>What exists already?</td>
</tr>
<tr>
<td>Methodology</td>
<td>How to achieve goals?</td>
</tr>
<tr>
<td>Results</td>
<td>What was achieved?</td>
</tr>
<tr>
<td>Discussion</td>
<td>How well were goals achieved?</td>
</tr>
<tr>
<td>Conclusions</td>
<td>What are the limitations?</td>
</tr>
<tr>
<td></td>
<td>Summary of achievements</td>
</tr>
<tr>
<td></td>
<td>What next?</td>
</tr>
</tbody>
</table>
Thesis Template

• Link: https://cs.ut.ee/en/content/thesis-deadlines-and-guidelines

• The typical structure of a thesis consists of the following components:
  • Title page
  • Information sheet
  • Table of Contents
  • Introduction
  • Terms and notions (optional)
  • Chapters
  • Summary
  • References
  • Appendices (if relevant)
  • License

Guidelines Document

<table>
<thead>
<tr>
<th>Thesis Components</th>
<th>Main Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>Why do it? (Context / Motivation)</td>
</tr>
<tr>
<td>Background</td>
<td>What exists already?</td>
</tr>
<tr>
<td>Methodology</td>
<td>How to achieve goals?</td>
</tr>
<tr>
<td>Results</td>
<td>What was achieved?</td>
</tr>
<tr>
<td>Discussion</td>
<td>How well were goals achieved?</td>
</tr>
<tr>
<td>Conclusions</td>
<td>Summary of achievements</td>
</tr>
<tr>
<td></td>
<td>What next?</td>
</tr>
</tbody>
</table>
Types of Theses


• Theoretical Study
  • Independent Study
  • Review-type Study

• Applied Research
  • Thesis based on a software solution created by the author
  • Hardware or software testing
  • Thesis based on study aid or learning material compiled by the author
Types of Theses


• Theoretical Study
  • Independent Study
  • Review-type Study

• Applied Research
  • Thesis based on a software solution created by the author
  • Hardware or software testing
  • Thesis based on study aid or learning material compiled by the author
Types of Theses – Theoretical


• Theoretical Study
  • Independent Study
  • Review-type Study

• Applied Research
  • Thesis based on a software solution created by the author
  • Hardware or software testing
  • Thesis based on study aid or learning material compiled by the author

More relevant for theoretical informatics
In SE/IS: Modeling, Algorithms, ??
Types of Theses – Theoretical


• Theoretical Study
  • Independent Study
  • Review-type Study
  • (Systematic) Literature Review
  • Mapping Study (=light-weight SLR)
  • Survey Study (questionnaires/interviews)

• Applied Research
  • Thesis based on a software solution created by the author
  • Hardware or software testing
  • Thesis based on study aid or learning material compiled by the author
Types of Theses – Applied


• Theoretical Study
  • Independent Study
  • Review-type Study

• Applied Research
  • Thesis based on a software solution created by the author
  • Hardware or software testing
  • Thesis based on study aid or learning material compiled by the author

*Design Science / Engineering / Case Study / Action Research*
Types of Theses – Applied


• Theoretical Study
  • Independent Study
  • Review-type Study

• Applied Research
  • Thesis based on a software solution created by the author
  • Hardware or software testing
  • Thesis based on study aid or learning material compiled by the author

Evaluation / Quality Assessment (Test & Static Analysis)
(and Process Assessment/Evaluation)
Types of Theses – Applied


• Theoretical Study
  • Independent Study  *Didactics / Evaluation based on Student Feedback*
  • Review-type Study

• Applied Research
  • Thesis based on a software solution created by the author
  • Hardware or software testing
  • Thesis based on study aid or learning material compiled by the author
Types of Theses – ACM SIGSOFT Standards

Action Research

*Empirical research that investigates how an intervention, like the introduction of a method or tool, affects a real-life context*

**Application**
- investigates a primarily social phenomenon within its real-life, organizational context
- intervenes in the real-life context (otherwise see the Case Study Standard)
- the change and its observation are an integral part of addressing the research question and contribute to research

**Important Elements (Attributes)**
- justifies the selection of the site(s) that was(we) studied
- describes the site(s) in rich detail
- describes the relationship between the researcher and the host organization
- describes the intervention(s) in detail
- describes how interventions were determined (e.g. by management, researchers, or a participative/co-determination process)
- explains how the interventions are evaluated
- describes the longitudinal dimension of the research design (including the length of the study)
- describes the interactions between researcher(s) and host organization(s)
- explains research cycles or phases, if any, and their relationship to the intervention(s)
- presents a clear chain of evidence from observations to findings
- reports participant or stakeholder reactions to interventions
- reports lessons learned by the organization
- researchers reflect on their own possible biases
Case Study

An empirical inquiry that investigates a contemporary phenomenon (the "case") in depth and within its real-world context, especially when the boundaries between phenomenon and context are unclear

Application

• Presents a detailed account of a specific instance of a phenomenon at a site. The phenomenon can be virtually anything of interest (e.g., Unix, cohesion metrics, communication issues). The site can be a community, an organization, a team, a person, a process, an internet platform, etc.

• Features direct or indirect observation (e.g., interviews, focus groups)

• Is not an experience report or a series of shallow inquiries at many different sites.

Important Elements (Attributes)

• Justifies the selection of the case(s) or site(s) that was(were) studied

• Describes the site(s) in rich detail

• Reports the type of case study

• Describes data sources (e.g. participants' demographics and work roles)

• Defines unit(s) of analysis or observation

• Presents a clear chain of evidence from observations to findings

• Provides supplemental materials such as interview guide(s), coding schemes, coding examples, decision rules, or extended chain-of-evidence tables

• Triangulates across data sources, informants or researchers

• Cross-checks interviewee statements (e.g. against direct observation or archival records)

• Validates results using member checking, dialogical interviewing, feedback from non-participant practitioners or research audits of coding by advisors or other researchers

• Describes external events and other factors that may have affected the case or site

• Uses quotations to illustrate findings

• Either: evaluates an a priori theory (or model, framework, taxonomy, etc.) using deductive coding with an a priori coding scheme based on the prior theory - OR: synthesizes results into a new, mature, fully-developed and clearly articulated theory (or model, etc.) using some form of inductive coding (coding scheme generated from data)

• Researchers reflect on their own possible biases
Data Science (exploratory)

Studies that analyze software engineering phenomena or artifacts using data-centric analysis methods such as machine learning or other computational intelligence approaches as well as search-based approaches.

Application
- Applies to studies that primarily analyze existing software phenomena using predictive, preemptive or corrective modelling.

Important Elements (Attributes)
- explains how and why the data was selected
- presents the experimental setup (e.g., using a dataflow diagram)
- describes the feature engineering approaches and transformations that were applied
- explains how the data was pre-processed, filtered, and categorized
- EITHER: discusses state-of-art baselines (and their strengths, weaknesses and limitations) - OR: explains why no state-of-art baselines exist - OR: provides compelling argument that direct comparisons are impractical
- defines the modeling approach(es) used (e.g. clustering then decision tree learning), typically using pseudocode
- discusses the hardware and software infrastructure used
- justifies all statistics and (automated or manual) heuristics used
- describes and justifies the evaluation metrics used
- goes beyond single-dimensional summaries of performance (e.g., average; median) to include measures of variation, confidence, or other distributional information
- discusses technical assumptions and threats to validity that are specific to data science
- provides a replication package including source code and data set(s), or if data cannot be shared, synthetic data to illustrate the use of the algorithms
- data is processed by multiple learners, of different types
- data is processed multiple times with different, randomly selected, training/test examples; the results of which are compared via significance tests and effect size tests (e.g., cross-validation)
- carefully selects the hyperparameters that control the data miners (e.g., via analysis of settings in related work or some automatic hyperparameter optimizer such as grid search)
- manually inspects some non-trivial portion of the data (i.e., data sanity checks)
Engineering Research (Design Science)

*Research that invents and evaluates technological artifacts*

**Application**
- Applies to manuscripts that propose and evaluate technological artifacts, including algorithms, models, languages, methods, systems, tools, and other computer-based technologies.

**Important Elements (Attributes)**
- describes the proposed artifact in adequate detail
- justifies the need for, usefulness of, or relevance of the proposed artifact
- conceptually evaluates the proposed artifact; discusses its strengths, weaknesses and limitations
- clearly indicates which (if any) empirical methodology is used
- EITHER: discusses state-of-art alternatives (and their strengths, weaknesses and limitations) - OR: explains why no state-of-art alternatives exist - OR: provides compelling argument that direct comparisons are impractical
- EITHER: empirically compares the artifact to one or more state-of-the-art alternatives - OR: empirically compares the artifact to one or more state-of-the-art benchmarks - OR: provides a clear and convincing rationale for why comparative evaluation is impractical
- assumptions (if any) are explicit, plausible and do not contradict each other or the contribution’s goals
- provides supplementary materials including source code (if the artifact is software) or a comprehensive description of the artifact (if not software), and any input datasets (if applicable)
- includes one or more running examples to elucidate the artifact
- evaluates the artifact in an industry-relevant context (e.g., widely used open-source projects, professional programmers)
Questionnaire Survey

A study in which a sample of respondents answer a series of (mostly structured) questions, typically through a computerized or paper form

Application

• Applies to studies in which:
  • a sample of participants answer predefined, mostly closed-ended questions (typically online or on paper)
  • researchers systematically analyze participants’ answers

Important Elements (Attributes)

• identifies the target population and defines the sampling strategy (see the Sampling Supplement)
• describes how the questionnaire instrument was created
• describes how participants were selected or recruited (e.g., sampling frame, advertising, invitations, incentives)
• step-by-step, systematic, replicable description of data collection and analysis
• describes how responses were managed/monitored, including contingency actions for non-responses and drop-outs
• EITHER: measures constructs using (or adapting) validated scales - OR: analyzes construct validity (including content, convergent, discriminant and predictive validity) ex post
• explains handling of missing data (e.g., imputation, weighting adjustments, discarding)
• analyzes response rates
• acknowledges generalizability threats; discusses how respondents may differ from target population
• provides the questionnaire instrument (as an appendix or supplementary materials)
• the questionnaire design matches the research aims and the target population
• provides supplementary materials including instrument(s), code books, analysis scripts and dataset(s)
• characterizes the target population including demographic information (e.g., culture, knowledge)
• accounts for the principles of research ethics (e.g., informed consent, re-identification risk)
• explains and justifies instrument design and choice of scales (e.g., by research objectives or by analogy to similar studies)
• validates whether the instrument's items, layout, duration, and technology are appropriate (e.g., using pilots, test-retest, or expert and non-expert reviews)
Systematic Reviews
A study that appraises, analyses, and synthesizes primary or secondary literature to provide a complete, exhaustive summary of current evidence regarding one or more specific topics or research questions

Application
• Applies to studies that systematically find and analyze existing literature about a specified topic
• Applies both to secondary and tertiary studies

Important Elements (Attributes)
• identifies type of review (e.g., scoping review, meta-analysis, systematic mapping study, etc.)
• presents step-by-step, systematic, replicable description of search process including search terms
• defines clear inclusion and exclusion criteria
• specifies the data extracted from each primary study; explains relationships to research questions
• describes in detail how data were extracted and synthesized (can be qualitative or quantitative)
• describes coding scheme(s) and their use
• clear chain of evidence from the extracted data to the answers to the research question(s)
• presents conclusions or recommendations for practitioners/non-specialists
• provides supplementary materials such as protocol, search terms, search results, selection process results; complete dataset, analysis scripts; coding scheme, examples of coding, decision rules, descriptions of edge cases
• mitigates sampling bias and publication bias, using some (not all) of:
  • (i) manual and keyword automated searches;
  • (ii) backward and forward snowballing searches;
  • (iii) checking profiles of prolific authors in the area;
  • (iv) searching both formal databases (e.g., ACM Digital Library) and indexes (e.g., Google Scholar);
  • (v) searching for relevant dissertations;
  • (vi) searching pre-print servers (e.g., arXiv);
  • (iiv) soliciting unpublished manuscripts through appropriate listservs or social media;
  • (iiiiv) contacting known authors in the area
Goal Setting & Baselining

• It is important to have one (or more) goal(s) set out at the beginning of the thesis
  • Question to ask yourself: What will you have achieved when you submit?

• There are many types of goals (see next slide)
  • Type of goal depends on type of thesis and problem statement
  • Goals might be formulated as Research Questions (RQs)

• In order to be able to decide whether you have achieved the goal(s), you need to know what the starting point is -> Baselining
Types of RQs

Research Question
- Exploratory Question
- Design Question
- Knowledge Question
- Causality Question

Existence Question
- Description and Classification Question
- Descriptive Comparative Question

Base-Rate Question
- Frequency and Distribution Question
- Descriptive-Process Question
- Relationship Question

Causality Question
- Simple Causality Question
- Causality-Comparative Question
- Causality-Comparative Interaction Question
Exploratory Questions

• **Existence questions** -> Does X exist?
  – Example: *Do code smells that trigger refactoring exist?*

• **Description and classification questions** -> What is X like? / What are its properties? / How can it be categorized? / How can we measure it? / What is its purpose? / What are its components? / How do the components relate to each other?
  – Example: *What are all the types code smells in Java code?*

• **Descriptive comparative questions** -> How does X differ from Y?
  – Example: *How do Java code smells differ from Python code smells?*
Knowledge and Design Questions

• **Knowledge Questions**: focusing on the way the world is
  • Questions about the normal pattern of occurrence of a phenomenon *(Base-rate Questions)*
  • Questions about relationships between two different phenomena *(Relationship Questions)*
  • Questions about causality between two phenomena *(Causality Questions)*

• **Design Questions**: concerned with how to do things better
Knowledge Questions

• Base-rate:
  • Frequency and Distribution Questions -> How often does X occur? / What is an average amount of X?
    Example: How many distinct code smell types do in Java code exist? How often do they occur?
  • Descriptive-Process Questions -> How does X normally work? / What is the process by which X happens? / In what sequence do the events of X occur?
    Example: How do software developers at company XYZ test their web-applications? (i.e., what processes and tools do they use?)
Knowledge Questions (cont’d)

- Relationship:
  - Relationship Questions -> Are X and Y related? / Do occurrences of X correlate with occurrences of Y?

Example 1: Do project managers’ claims about the intensity of testing correlate with the actual intensity of testing?
Example 2: Does the occurrence of certain types of failures correlate with specific use cases?
Knowledge Questions (cont’d)

• Causality:
  • Simple Causality Questions -> Does X cause Y? / Does X prevent Y? / What causes Y? / What are all the factors that cause Y? / What effect does X have on Y?
    Example: Does the use of UML diagrams improve the quality of the design?
  • Causality-Comparative Questions -> Does X cause more Y than does Z? / Is X better at preventing Y than Z?
    Example: Does the use of UML diagrams improve the quality of the design more than other graphical design notations?
  • Causality-Comparative Interaction Questions -> Does X or Z cause more Y under one condition but not others?
    Example: Does the use of UML diagrams improve the quality of the design more than other graphical design notations in large projects, but not otherwise?
Design Questions

-> "What is an effective way to achieve X?" / What strategies help to achieve X?"

Examples:

What is an effective way for teams to capture requirements in order to improve communication with customers?

or

What is an effective way for developers to represent design knowledge in order to improve design quality?
Methods -> Literature

• ACM SIGSOFT Standards of (Empirical) SE Research:

• Systematic Literature Survey:

• Mapping Study:

• Case Study:

• Design Science:
Milestones

• 19-Sep-2022: MSc topic and supervisor confirmed
  • e-mail from supervisor to dietmar.pfahl@ut.ee required
• 21-Sep-2022: Last possibility to withdraw from the seminar (without going through the Dean's office)
• 26-Sep-2022: Classification of thesis topic
• 17-Oct-2022: 1st draft of report
• 14-Nov-2022: 2nd draft of report + presentation slides (for presenters on 16-Nov-2022)
• 21-Nov-2022: 2nd draft of report + presentation slides (for presenters on 23-Nov-2022)
• 28-Nov-2022: 2nd draft of report + presentation slides (for presenters on 30-Nov-2022)
• 05-Dec-2022: 2nd draft of report + presentation slides (for presenters on 07-Dec-2022)
• 12-Dec-2022: 2nd draft of report + presentation slides (for presenters on 14-Dec-2022)
• 19-Dec-2022: Final report submission
Schedule — mandatory sessions in red

- **07-Sep-2022: Introduction Session** - look at requirements for graduation theses at ICS
  - In this session, I give an overview of the seminar course → Slides
- **14-Sep-2022: Consultation Session** (please reserve individual time slot during first 15 min of session) - first 15 min will be used for general Q/A
  - In this session, I answer questions related to the first submission (classification of topic according to the guidelines for graduation theses) and other questions about the seminar and its next submissions.
- **No formal sessions from 21-Sep-2022 to 09-Nov-2022:**
  - Please schedule individual consulting sessions when needed. Preferably, I try to schedule individual consultation sessions on Wednesdays after 16:15. If that does not work for you, we can try to find a different time-slot. Please contact me for an individual consultation at least 24 hours ahead of time.
  - Different to all other sessions, consultations will be done online (use Zoom link in Moodle)
- **16-Nov-2022: Presentation Session 1**
- **23-Nov-2022: Presentation Session 2**
- **30-Nov-2022: Presentation Session 3**
- **07-Dec-2022: Presentation Session 4**
- **14-Dec-2022: Presentation Session 5**
Assessment and Grading

• This seminar is a "pass/fail" course.

• Assessment criteria of the seminar
  • Presence during mandatory sessions (if online session, then video must be switched on)
  • Punctual submission of information / documents
  • Quality of final report -> Grades 1-5 possible
  • Quality of presentation (incl. slides) -> Grades 1-5 possible
    • Quality characteristics: (1) Use of time; (2) Slides (text and visuals); (3) Content; (4) Presentation
  • Giving feedback to at least three other students during/after presentation sessions
    • Quality characteristics: (1) Use of time; (2) Slides (text and visuals); (3) Content; (4) Presentation
  • See scoring rubric on next slide
# Presentation Scoring Rubric

(Rate each category 0-5, with 5 being the highest possible score)

<table>
<thead>
<tr>
<th>Item being evaluated</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Points earned</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use of time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>More than 50% time overrun or less than 50% of time used</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slides (text and visuals)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Letters too small to read and slides too full with text / no visuals or visuals are misleading and overly complex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Content</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No goals stated + no motivation + no method description + no related work/baseline</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presentation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presenter needs to read from slides / the flow of the presentation is lost several times / presenter cannot be heard</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Point reductions if an item is missing or incomplete

Point reductions if any item is missing or incomplete


Large enough letters (1) / Slides not overfull with text [2] / visuals support content [2]


Presenter seems to understand the content and uses the slides as support (no reading from slides) [3] / Clearly understandable voice [2]

Total | ? /20
Assessment and Grading (cont’d)

Grading Scheme for Presentation & Final Report

- Grade 1 (very poor) should be used only in exceptional cases: if the thesis clearly violates the basic principles of an academic work (for example, in the case of plagiarism). Grade 1 in any of the criteria automatically means F (fail) as a final grade.

- Grade 2 (poor) is a negative grade which should be used if the thesis does not meet the minimal requirements established to the criterion concerned. Grade 2 in two criteria automatically means F (fail) as a final grade.

- Grade 3 (barely acceptable) is the lowest positive grade which should be used in the case when the work meets the minimal requirements of the criterion, but contains some substantial shortcomings.

- Grade 4 (good) means that the thesis is a good work without major shortcomings.

- Grade 5 (very good) requires that the work is almost flawless according to the given criterion and is outstanding in some aspect.
Assessment and Grading (cont’d)

How to **definitely** fail the course?

- Not being present in **all** mandatory sessions (without sufficient justification; sufficient justification is a certificate from a medical doctor saying that you cannot be at work).
- Missing at least 2 milestones by more than 24 hours - no further exceptions will be granted!
- Not submitting **all** requested information items and documents (topics, supervisors, reports, slides, feedback).
- Not giving a presentation
- Not giving feedback to at least three other students' presentations (for a session in which you do not present yourself)
- Having a grade below '3' in the final report or the presentation
Communication

• Use Slack

SE-Sem2-2022.slack.com
Hints & Tips – Technical Writing

Importance of good writing

• It helps your readers to:
  • distinguish between more important and less important material
  • quickly find specific information
  • grasp the flow of argument, the assumptions (incl. previous/related work), and the context
• Remember: Often you don’t have control over the recipients of your document
Hints & Tips – Technical Writing

Style issues:
• No colloquial expressions or slang!
• Appropriate usage of terminology
• Consistent wording (one concept – one expression)
• Short sentences
• Normally: result-oriented (not process-oriented)
• Introduce abbreviations before first usage
• Neutral style instead of “we …”, “our …”, etc.
• After each headline there should be some text …
Figures and tables:

• Have a caption (incl. brief description and numbering)

• Are referenced from within the related text

density? Each question is then refined into measures that can be collected on the field. For instance, defect density may be defined as the ratio of the number of defects found to the number of lines of code. Figure 1 shows this top-down refinement of goals into measures. Figure 1 also shows that several measurement goals may be pursued at the same time, and questions and measures may be reused across goals, thus decreasing the effort for adding further goals to an existing set of goals, questions, and measures.
Hints & Tips – Technical Writing

References:

• All literature included in the list of references must be referenced (at least once)

References:


those that refer to the way software relates to its development or operational environment. Process attributes are discussed in Section 9. Remarks on the practical application of software measurement are in Section 10. Possible future developments are discussed in Section 11.

Good surveys of the state of the art and on-going research can be found in [1, 2].
How bibliography needs to be organized? References should appear as …

References should appear as …

• **Book with one author:**

• **Book with two authors:**

• **Book with more than two authors:**

• **Journal article:**

• **A publication in press:**
How bibliography needs to be organized?
References should appear as ...

References should appear as ...

- **Journal article:**

- **A publication in press:**

- **Report from a university:**

- **Published proceedings:**
How bibliography needs to be organized?

References should appear as ...

References should appear as ...

- **Unpublished doctoral dissertation or master’s thesis:**

- **A presented paper:**

- **Web site:**
In-Text Citations

• In-text citations should appear with author surname followed by publication year in parentheses
  (Brown, 2002)

• Citing several references in-text:
  In most organizations, data resources are considered to be a major resource (Brown, 2002; Krall & Johnson, 2005; Smith, 2001).
  Brown (2002) states that the value of data is recognized by most organizations
  “In most organizations, data resources are considered to be a major organization asset” (Smith, 2001, pp. 35-36) and must be carefully monitored by the senior management.
  Brown (2002) states that “the value of data is realized by most organizations” (p. 45).

• If you have organized the citations with number in brackets:
  In most organizations, data resources are considered to be a major resource [15, 30, 84].
Requirements for Thesis Draft

• Templates and Guidelines: https://cs.ut.ee/en/content/thesis-deadlines-and-guidelines

• The **final draft (in total around 12-15 pages) should contain the following**
  • Title page
  • Abstract -> results part and conclusions can be empty
  • Table of Contents
  • Introduction / motivation / problem statement (section **Introduction**) -> 1-2 pages
  • Literature review (section **Related Work**) -> preliminary summary of core literature -> 2-5 pages;
  • Baseline (section **Background**) -> if you have a design science thesis: must contain baseline -> 2-3 pages
  • Plan of the contribution (section **Method**) -> must be detailed enough to understand what exactly will be done to achieve the goals -> 4-7 pages
  • Intermediate results (section **Results**) -> can be empty
  • Discussion of results / limitations (section **Discussion**) -> can be empty
  • Summary and conclusions (section **Conclusion**) -> can be empty
  • List of references -> at least 5 references expected (must be well-formed)
Evaluation of Thesis Drafts

- I will grade your drafts using the evaluation criteria 1-3 available at
  - *See Section 5.1: Content / Level of Complexity / Appearance*

- Additional rules
  - Draft contains a copy/paste of more than 10 consecutive words from any source → -1 point per copied fragment
  - Reproducing/re-drawing 1 or 2 pictures or tables is tolerated, if you explicitly cite the origin of the picture/table in the caption, else -1 point
Tips for Writing a Thesis

How theses get written: some cool tips ...

Hints & Tips – Presentation

• What to present:
  • Problem Statement (Motivation)
  • Goals (RQs)
  • Baseline
    • Related work done by others or
    • Current state-of-practice (if thesis done in a company or building on top of existing materials)
  • Method
    • Steps how you plan to achieve the goals (construction / evaluation)
    • Already achieved results (if any)
  • Plan of completion (schedule until submission)
Hints & Tips – Presentation

• Evaluation criteria (up to 1 point per criterion):
  • **Slides**: Is the amount of text appropriate? Are figures and tables used appropriately where possible?
  • **Public speaking/oratory**: Is the posture appropriate? Is the voice level and intonation engaging? Is the gesturing appropriate? Is the rhythm engaging?
  • **Structure**: Is the structure of the presentation clear? Are the transitions between logical parts of the presentation clearly articulated?
  • **Content**: Does the presentation highlight the main points of the paper? Is the level of detail appropriate? Are examples used appropriately?
  • **Timeliness**: Does the presentation use the allocated time appropriately? Is the time limit respected?
Hints & Tips – Presentation

• How to present:
  • Use the time – but don’t exceed (12 min presentation + 3 min discussion)
  • Balance content (per slide and across whole presentation)
  • Speak clearly (not too fast / not too low)
  • Let the slides support your talk – don’t use slides to write down what you say
  • Use visuals/diagrams/tables where appropriate
How **not** to do a presentation ...

- Don McMillan: Life After Death by PowerPoint
  - [http://www.youtube.com/watch?v=WGiePuNFXwY](http://www.youtube.com/watch?v=WGiePuNFXwY)
  - [http://www.youtube.com/watch?v=zDvm1PVtgWo](http://www.youtube.com/watch?v=zDvm1PVtgWo)
  - [http://www.youtube.com/watch?v=lpvgfmEU2Ck](http://www.youtube.com/watch?v=lpvgfmEU2Ck)