Research Ethics and Scientific Publication: The Key to Success

Professor Guttorm Sindre
Structure of the talk

1. Two types of theses
   • And some simple advice related to planning and writing
2. Where to publish papers?
3. How to get accepted?
4. Research and publication ethics
5. Literature review / survey papers
To get a phd, results must be

- Original
  - Not just for you, but for the world
- Important
  - For academia and/or for industry
- Correct
  - Technically, methodologically, and ethically sound
- Convincingly presented
  - Monograph thesis OR
  - Paper collection thesis (intro/synthesis + N papers)
Contributions

• For a paper: One key contribution enough
• For a thesis:
  • Normally need several contributions
    • Unless extremely original and important one
  • Contributions should be closely related
    • Fit inside one thesis topic
    • This applies regardless of thesis type
• Different kinds of contributions
  • Contributions to theory
  • Contributions to practice
  • Contributions to method

<table>
<thead>
<tr>
<th></th>
<th>P1</th>
<th>P2</th>
<th>P3</th>
<th>P4</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C2</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>C3</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>C4</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>
Two types, pros and cons

• Monograph
  – PROS:
    • Easier to accommodate deep arguments and lots of detail
    • Thesis becomes a publication in its own right, separate from papers
    • Better structure, more pleasing to read?
  – CONS:
    • Less synergy between thesis writing and paper writing
    • Need more discipline to control the progress
  – RISK:
    • Early parts may become dated, especially if you lag behind

• Paper collection
  – PROS:
    • Dividing the work into smaller chunks, clear milestones
    • Stronger synergy between paper writing and thesis writing
    • Paper publication gives early external feedback
  – CONS:
    • Length restrictions for papers prevent long / deep sections
    • Overlap between papers
    • Gaps between papers?
  – RISK:
    • What if you do not get enough papers accepted?
    • Problems with synthesis?
Some simple advice

• On choice of thesis type:
  – discuss this with your supervisors, BUT
  – If you choose a monograph
    • Should still publish papers in parallel
  – If you choose a paper collection
    • Still need a structure for thesis as a whole

• Plan delivery 3 months before funding ends
  – Calculate backwards from this when making your plan
  – When will different contributions be readily researched?
  – When will papers be submitted, thesis chapters written?
Advice (cont.)

• Start writing as early as possible! (papers, thesis)
• Don’t be shy to show early drafts to supervisors
  – And others, e.g. fellow phd students
• Don’t be reluctant to publish early stage ideas if possible

• But… don’t continue publishing just early stage ideas all way through your phd studies
  – Papers must gradually build up to mature results
  – Gradually target more prestigious outlets
Advice, cont.

• Avoid becoming your own worst enemy
  – Inflated ego / arrogance / stubbornness / sloppiness
  – Lack of focus
  – Lack of self-confidence
    • Fear of writing / putting your ideas on paper
    • Fear of proposing ideas or showing written material to others
  – Exaggerated self-criticism / perfectionism
    • too much time polishing one chapter or paper
    • other tasks remain pending

• Good enough is better than nothing
  – Your most brilliant work can come during a postdoc or later
Where to publish papers?

- **Different publication outlets:**
  - **Doctoral consortia:** Research plans
    - NOTE: normally, these papers cannot be part of the thesis collection
  - **Workshops:** Early stage ideas, with limited validation
    - Often narrower focus than conferences
    - NOTE: workshop papers may be dubious as part of thesis
  - **Conferences:** Clear contribution, at least some validation
    - Different paper types:
      - full (~5000 words), short (~2500 words), posters (~1-2 p)
    - NOTE: normally only full papers as part of thesis
  - **Journals:** Mature work with detailed validation / survey papers
    - Allowing longer papers than conferences

- **Avoid bogus and poor quality outlets!**
What determines the prestige of publication channels?

• Should be peer reviewed
  – Other kinds of dissemination may also be valuable
    • E.g., industry forums
    • Papers in such casual channels not plausible as part of thesis
  – Never try to pass a paper off as peer-reviewed if it is not

• Journals
  – Circulation, #subscribers / readers
  – Professional opinion (ranking)
  – Acceptance ratio
  – Impact factor (IF)

• Conferences
  – Circulation of the proceedings, size (#attendants)
  – Professional opinion (ranking)
  – Acceptance ratio
  – Impact factor

• Acceptance ratios of conferences and journals not directly comparable
Publications and economy

- University departments get points for their publications
  - "Journal" paper: 3, 1 or 0 points
  - "Book section": 1, 0.7 or 0 points
  - See http://dbh.nsd.uib.no/kanaler/
  - Important to register all publications properly in Cristin
- Value of publication point has been gradually decreasing
  - Department gets approx 18 KNOK per point, 2 years later
- Before you submit, check funding
  - Are there costs related to the publication?
    - Conferences: registration fee, travel and accommodation
    - Journal: publication fee, page charges?
  - Will the department or project cover these costs?
  - Try to use alternative sources if possible
Submission & review processes

- Stage one: find a suitable publication outlet for your result
  - Journals
    - Submit any time, no deadline (except special issue)
    - Topic list, length restrictions
  - Conferences: **Call for papers**
    - Given deadlines (for submission and notification)
    - Topic list, paper length and format restrictions
    - Paper categories
    - Program committee etc.

- General advice: Read the topic list / CFP thoroughly
  - Is this the best place to send your work?
    - Good conference, appropriate topic list, reachable deadline, …
  - Which category of paper are you going for?
If uncertain…

…whether a publication channel has the appropriate research focus and quality level?

- Read some papers from previous issues of the journal or conference
- Check the topic list
- Check who is on the program committee or editorial board
- **Ask your supervisor**
How to get accepted (1)

• Some conferences have several paper categories
  – Example:
    • Technical solution papers
    • Scientific evaluation papers
    • Industrial experience papers
    • Theory / vision papers
  – Generally: full papers vs. short / r-i-p / position papers

• Advice:
  – Have a clear idea what type of paper you are trying to write
  – Be sure to satisfy the evaluation criteria for that category
How to get accepted (2)

• The research (and paper) must be
  – Original
    • Both vs. work by others and own previous publications
  – Significant: important findings, useful for somebody
  – Correct: e.g. proper use of methods, statistics, reasonable assumptions, …
  – Well presented (language, structure)
  – Relevant / within the scope of the particular conference or journal

• Key concerns
  – Make a precise claim
  – Validate the claim in an appropriate manner
  – Good comparison with related work
    • Helps to convince about originality (and maybe also significance)
  – Follow format and length constraints
  – Have someone look over the language if your English is not very good
    • Professional proof-reading services also a possibility, but cost money
Guidelines to successful papers

- Shaw: "Writing good software engineering research papers" (Proc. ICSE’03)
- CHI guide to successful papers
- Both target specific research communities
- But contain mostly generic advice
- If limited writing skills is an obstacle for you…
  - consider taking a course in Academic Writing
**IMRAD**

- Typically recommended structure for papers, especially in empirical disciplines
  - Introduction
  - Methods
  - Results, And
  - Discussion

- Sections need not have those exact names, and there could be other sections in addition, but this is an overall structure often seen
Publication / research ethics

• Unethical behavior **must be avoided**
  – Plagiarism
  – Un- / mis-declared assistance
  – Fabrication
  – Unauthorized handling of sensitive or confidential data
  – Research with harmful effects
  – ”Self-plagiarism” / improper reuse
  – Corruption / failure to declare conflict of interest

• For some types of research, application to an ethical committee would be necessary before collecting data
Plagiarism

• Usage of somebody else’s text / ideas / code / data / results WITHOUT properly crediting the originator
  – i.e., the reader may be misled into believing it is your own text / ideas / code / …

• Examples:
  – Direct copy-paste of text from somebody else, without quoting info
    • Or direct translation, without info about this
  – Minor rephrasing of the text, without reference
    • Or even with reference, but scope of reference is unclear
  – Direct transcript of oral statements, without crediting speaker
  – Reuse of somebody else’s code in own prototype with no mention
    • i.e., reader may think you coded more than you actually did
  – Unacknowledged reuse of somebody else’s data for a new analysis
    • i.e., reader may think you collected the data
  – Citing self rather than original source

• NOTICE: plagiarism ≠ copyright or license infringement
Avoid plagiarism — quotes and references

• All quotes must be specifically marked as such
  – "quoted text" (reference, page number)
  – For longer passages e.g. (begin quote) quoted text (end quote)
  – For quotes, always include page number(s) with the reference
    • Just the reference is not enough if the text is identical
    • Page numbers help reviewers check the quote quickly
  – Check that your quote is really the original source
    • Or you may risk quoting a plagiarizer
• Never paste other text directly into your own w/o quotes
  – Not even for internal documents "under work"
    • Risk of accidental plagiarism (forgets to rephrase, gets submitted)
    • Collaborations: one person does not know that it is copy-paste
• Rephrased text, no quote marks, but ensure a clear scope of reference
  • E.g., "This entire paragraph / subsection / bullet list is based on [5]…"
  • (not necessarily a good paper – maybe too little originality? – but at least you avoid plagiarism since the source is clearly given so you do not pretend that the ideas presented in that part of the manuscript are your own)
What if somebody plagiarizes you?

• Work you already published…
  – Possible actions:
    • Gather evidence (you published first, exact parts plagiarized)
    • Inform your superiors (e.g., supervisor, dept head)
    • Contact plagiarism publisher
      – Demand that article is withdrawn
      – Demand a note of apology?
    • Contact home institution of plagiarizing authors?
    • Contact copyright owner of the original publication?

• Work you did not publish yet
  – An even worse problem
    • especially if the material is essential for your thesis
  – Many of the same actions
  – Even more important to gather good evidence
Un/mis-declared assistance

• Somebody contributed to the work
  – But is not included in the author list, or mentioned in the text
  – Or: Did not contribute, but is included

• Specifically for scientific papers:
  – The Vancouver Convention presents criteria for what kinds of contributions qualify for a place in the author list

• To include joint papers in a thesis:
  – Need to declare amount and nature of own contribution
  – Signed statements from co-authors that this is correct
  – Advice: Do this as early as possible
    • not wait until thesis delivery
Fabrication

• “…the truth, the whole truth, and nothing but the truth”
  – Not lying, nor deliberately misleading the reader

• Examples of fabrication
  – Constructing fake data
    • E.g. from experiments, questionnaires, interviews, simulations, benchmarks
  – Editing data sets or results, e.g. to fit a hypothesis
    • E.g., removing ”outliers” without proper discussion
  – Biased design or misleading presentation of research procedures
  – Pretending to have done more than you really did
    • E.g., passing screen sketches off as a running prototype
  – Selective presentation or publishing of results
    • E.g., one experiment positive, another negative, publishing only one
  – Biased editing of quotes, biased presentation of related work

• Be explicit and precise about methods and results
  – And discuss threats to validity where relevant

• Avoid citing papers not read (or use ”[5] (as cited by [6]) …”)
Unauthorized handling of sensitive / confidential data

- Colloborating with companies?
  - May some data be confidential?
  - Beware what data you are allowed to see
    - and what you are allowed to publish

- Collecting data about persons?
  - May need to apply to ethical committees
    - BEFORE data collection begins
  - E.g., health data are always sensitive

- Avoid going into projects where you risk not to be allowed to publish your findings
  - Clarify such risks beforehand
  - Have agreements / permissions in writing where relevant
Research with harmful effects

• Either: The research process is harmful to somebody?
  – Animals? Environment?
  – Humans? (e.g., participants in experiments)
  – Resources / Infrastructure (e.g. faulty code crashing servers)

• Or: The end product of the research is harmful to somebody?
  – Often a weighting of harm vs. benefit

• Application to ethical committee could be needed
”Self-plagiarism”

• Trying to use the same or very similar work twice
  – E.g., using the same term paper in two different courses
  – Publishing same paper in several outlets

• Could be OK in some situations
  – But generally NOT
  – Most conferences / journals demand that works be original and neither published nor under review elsewhere

• If submitting an extended version of a conference paper to a book or journal…
  – Make it explicit (to the editor and in the text) that it is an extension, and what the extended content is

• Reusing background material from previous papers?
  – be careful of direct copy even here
  – especially if there is author list discrepancy
Corruption / unethical reviews

• As an author submitting a paper
  – Bribing or threatening a reviewer or editor
  – Trying to uncover reviewer identity (if blind review)
    • And taking revenge for rejections if possible
  – …or revealing own identity (if double-blind review)

• As a reviewer
  – Giving biased reviews (positive to friends, negative to enemies)
    • If anything prevents you from reviewing a paper objectively, declare conflict-of-interest
  – Using material from unpublished papers you are reviewing

• As author, reviewer or editor
  – Biased boosting of own or other citations
    • E.g., citation rings among authors
    • Editors demanding that papers cite other papers from same journal
Literature review and survey papers – what and why?

• Literature review as part of
  – thesis, for instance chapter(s) called "State-of-the-art" or "Related Work"
  – papers with own contributions, e.g. sections called "Background" or "Related Work"

• As a paper in its own right
  – does NOT present new ideas or findings by the author
  – surveys the state-of-the-art within a certain research area, e.g.
    • Available technology within some domain
    • Available algorithms or methods to do some task / solve some problem
    • Existing studies to measure effectiveness or user-satisfaction of some type of IT
  – Most often found in journals
The bad review

• Obvious mistakes
  – Overlooking or omitting relevant works
  – Including irrelevant works (e.g., nepotism)
  – Falsification, fabrication, plagiarism
  – Sloppiness (e.g. incorrect reference details) and poor writing

• Less obvious mistakes
  – List appearance, lack of synthesis
  – Poor prioritization (how much space to spend)
  – Just agreeing or disagreeing, no critical appraisal
  – Over-usage of direct quotations, rather than own argument
  – Opinions not substantiated
  – Only shallow comparisons
Avoid too much quoting anyway!

• Direct reproduction of others’ arguments
  – makes your text dull and less innovative
  – often gives a fragmented structure and poor flow of argument
  – If you quote large fragments, this could be breach of copyright, even if proper quoting means that plagiarism is avoided

• The same goes for minor rephrasings which are close to direct reproduction

• While proper use of ”as cited by” solves ethical issues of citing unread papers, it signals
  – lack of independent opinion
  – laziness (relying on others, not doing the work yourself)
Example

• Assume our text is
  – According to [1] requirements engineering (RE) can be defined as “(some definition)” (p 5), and is important for the following reasons: (some reasons). [1] considers RE to consist of the following tasks: (list of tasks), and sees (list of problems) as the main research challenges. [2] defines RE as “(some other definition)” (p 11), and stresses the distinction between various types of requirements: (list of types). The importance of a good specification is much stressed (list of arguments). [3] defines requirements engineering as “(definition)” (p 14) <this def. btw is equal to [1], which is cited in [3] at this point> and claims that this is vital to any project because (list of arguments) ... etc. (for source [4], [5], ...)

• Assume that there is proper referencing and quoting here, yet what is the problem?

• Better (for example)
  – There are several definitions of requirements engineering (RE), here the one by [1] is chosen: “(chosen definition)” [1] (p 5). Several arguments are commonly made for the importance of RE:
    • Argument 1... (references to sources making this argument, say [1,3])
    • Argument 2... (references to sources making this argument, say [1,2,4])
    • ...

Literature review / Survey papers

• Highly relevant for phd students
• Find good survey papers within your field
  – May greatly help your own literature search
  – and the writing of your state-of-the-art
• Write a survey paper yourself?
  – High synergy with the state-of-the-art
  – Good opportunity to get a journal publication
    • Already at an early stage
  – If successful, may get a lot of citations
  – BUT: Must give added value beyond existing survey papers
Summing up

- Start writing early (both papers and thesis)
- Don’t be afraid to show early drafts to your supervisor and others who could give good comments
- Investigate early about possible publication outlets in your research area
- Make sure you follow ethical standards throughout your research and publication activities, see also:
  - [https://innsida.ntnu.no/wiki/-/wiki/English/Ethics+at+NTNU](https://innsida.ntnu.no/wiki/-/wiki/English/Ethics+at+NTNU)
- Consider publishing a survey paper as a result of your literature review