POLYTECHNIQUE
MONTRÉAL

TECHNOLOGICAL
UNIVERSITY

INF[67]900E Lecture 4 – Writing Papers/Proposals

Dr. Bentley Oakes
bentleyjoakes.github.io

Deliverables

Sept 26th

- One page summary of a paper
- Hand-in the summary and the PDF on Moodle
- Evaluated on being **clear, specific, and structured**
- **Note: I'm at a conference all this week**

Oct 3rd

- Two page critical review of a paper
- Template and excellent examples on Moodle
- Evaluated on being **constructive**, **specific, professional, structured**

Oct 10th

- Evaluation of another student's review
- Performed through Moodle, few sentences per criteria above
- Evaluated on having **lessons present, lessons missing, general feedback**

Last Lecture

1. Review process
3. Review structure
4. Review principles

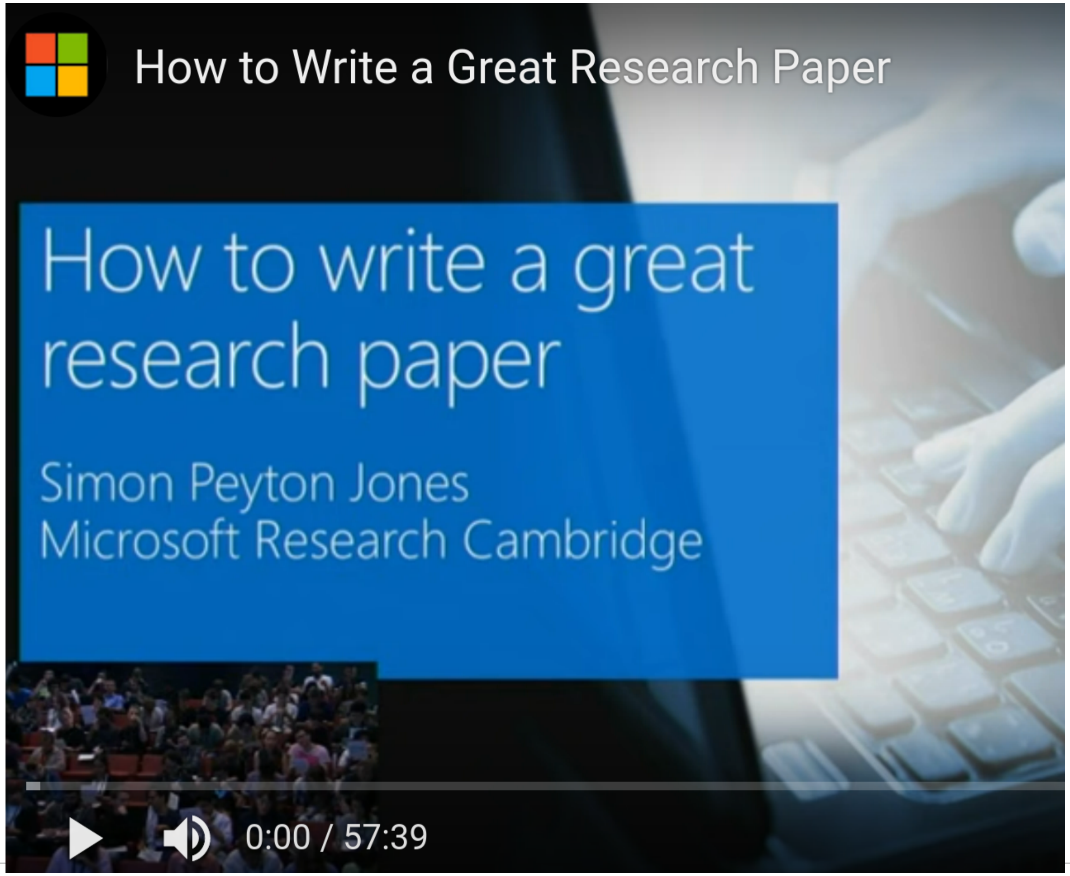
This Lecture

1. Plan your writing
2. Plan your research
3. Paper structure
4. Proposal structure

Required Reading/Watching

https://www.microsoft.com/en-us/research/academic-program/write-great-research-paper/

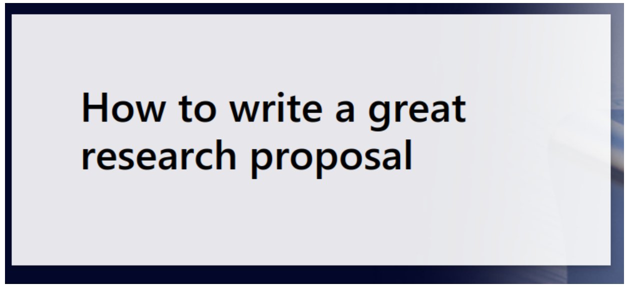
On Moodle



How to Write a Great Research Paper

How to write a great research paper

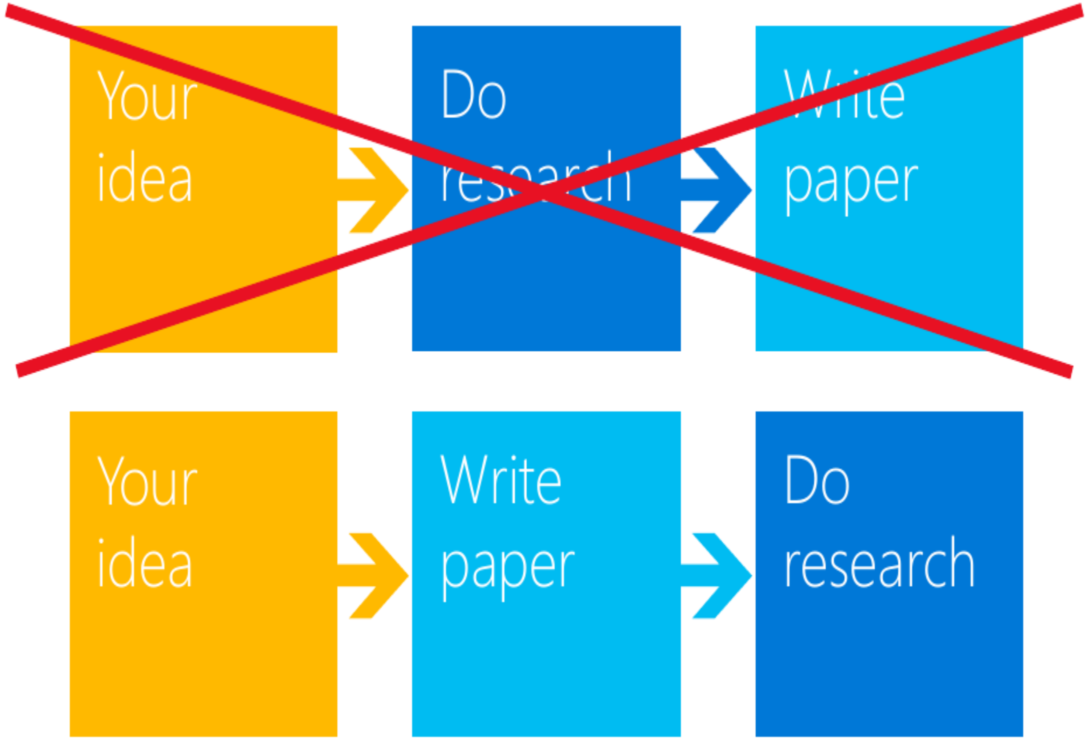
Simon Peyton Jones
Microsoft Research Cambridge



How to write a great research proposal

Plan Your Writing

Plan Your Writing



~~Your idea~~ → ~~Do research~~ → ~~Write paper~~

Your idea → Write paper → Do research

Paper writing:

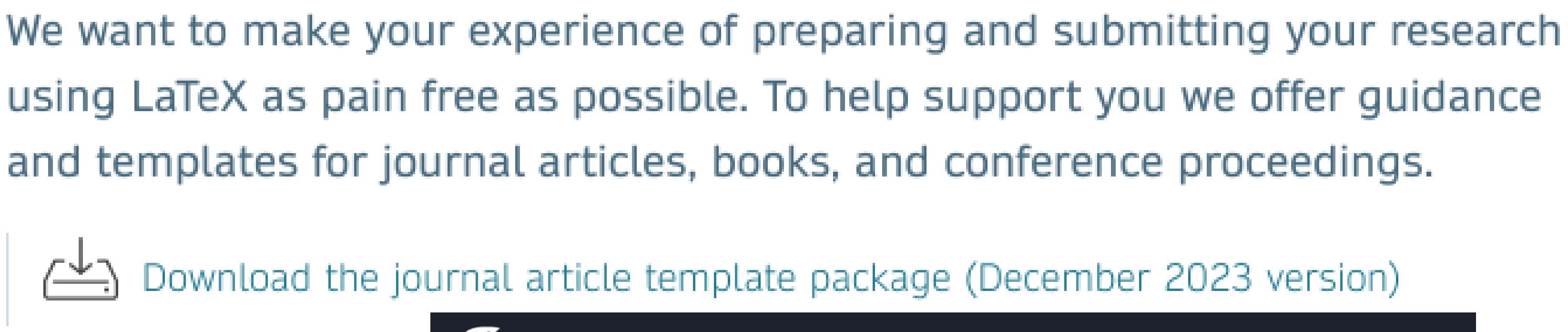
- Forces us to be clear, focused
- Crystallises what we don't understand
- Opens the way to dialogue with others: reality check, critique, and collaboration

Plan Your Writing

- What kind of document are you writing?
  - Paper, proposal, thesis
- What is the deadline?
  - Be deadline-driven
- What is the paper about? What are the contributions?
  - More on this later

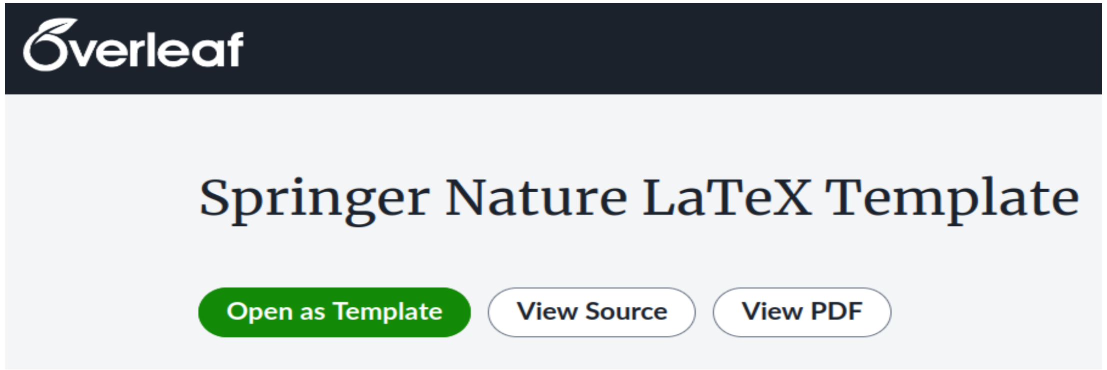
How to Get Started

- Find the right LaTeX format and structure



We want to make your experience of preparing and submitting your research using LaTeX as pain free as possible. To help support you we offer guidance and templates for journal articles, books, and conference proceedings.

Download the journal article template package (December 2023 version)



Springer Nature LaTeX Template

NSERC Online presentation and attachment standards

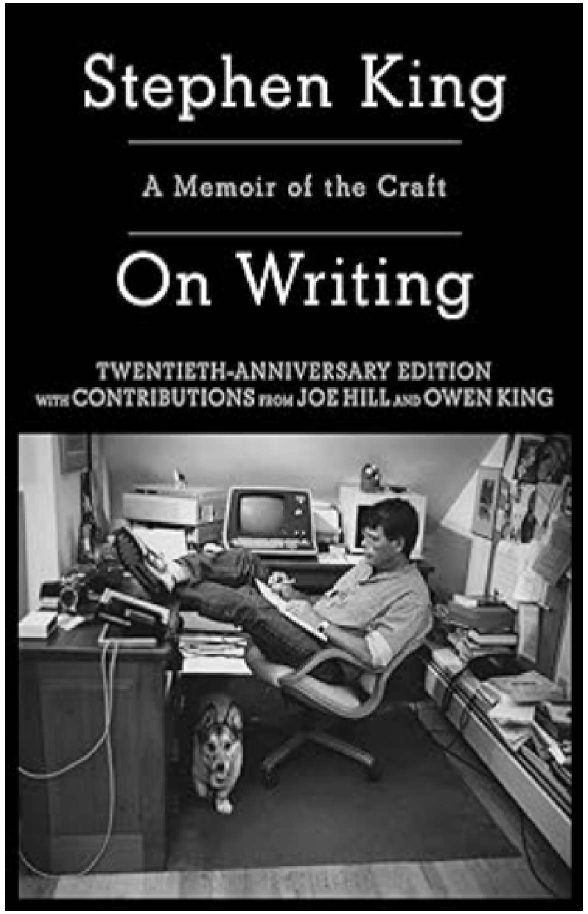
How to Get Started

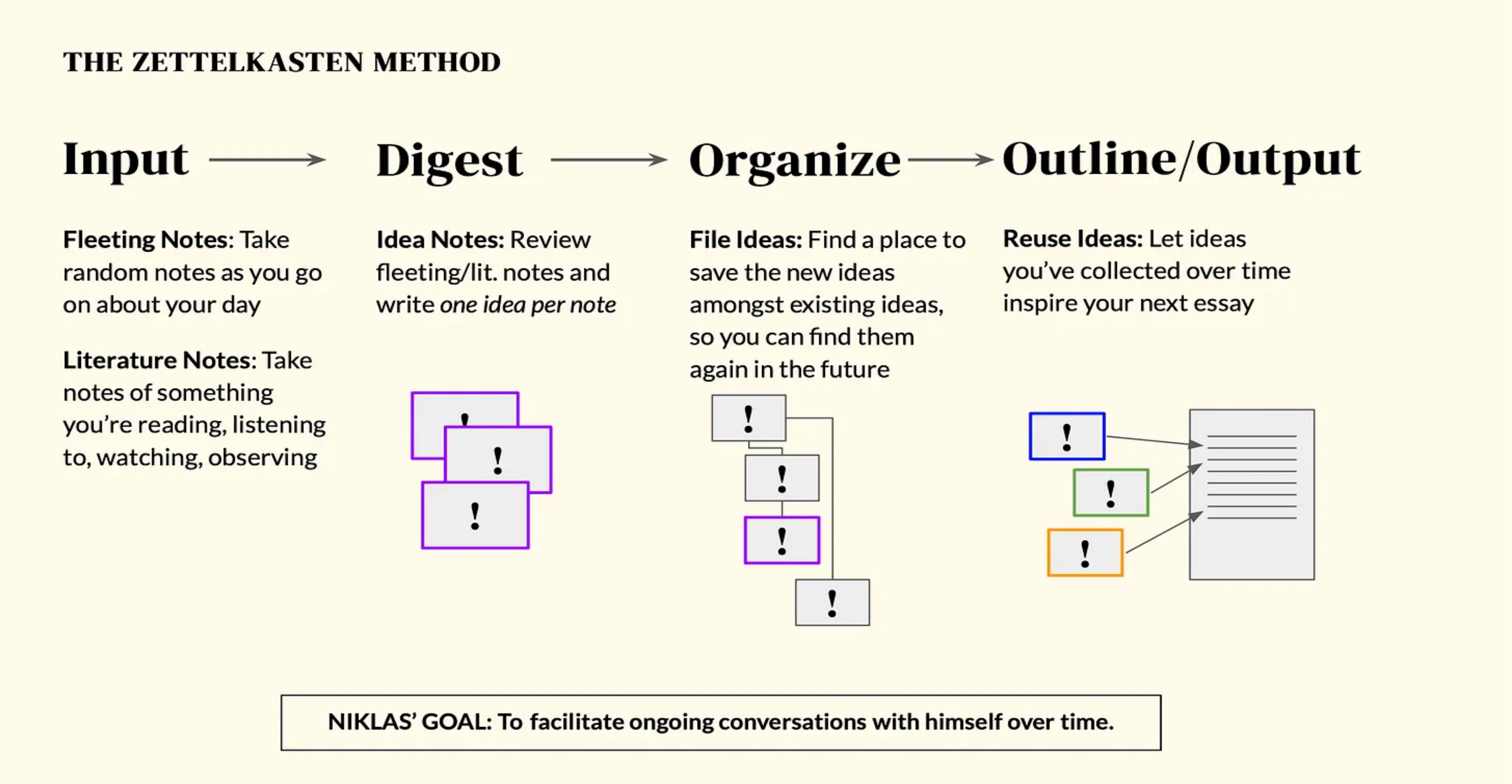
- Fill in the basic structure
  - Sections and sub-sections headers

- Write anything
  - Text is easy to change
  - Editing is much easier than writing!
  - Focus on ‘drafts’ for iteration
  - Iterate, iterate, iterate…

Come up with a schedule and a place for writing

- Writing is mentally taxing and shouldn’t be rushed


Stephen King
A Memoir of the Craft
On Writing
TWENTIETH-ANNIVERSARY EDITION
WITH CONTRIBUTIONS FROM JOE HILL AND OWEN KING



THE ZETTELKASTEN METHOD

Input → **Digest** → **Organize** → **Outline/Output**

Fleeting Notes: Take random notes as you go on about your day

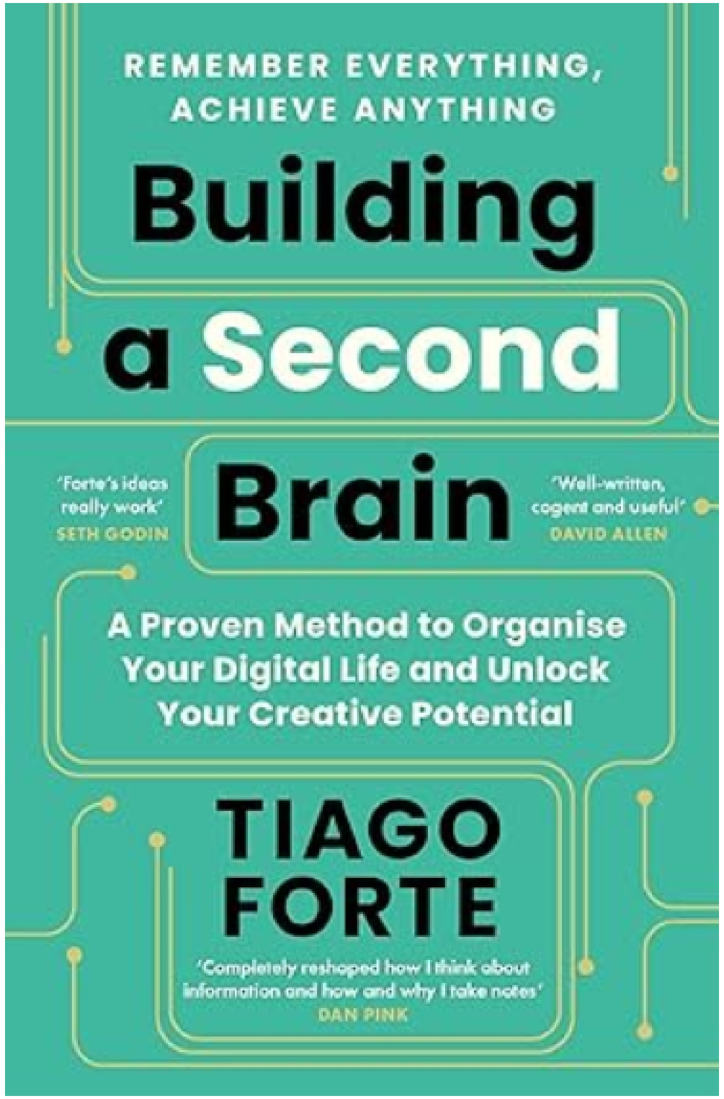
Literature Notes: Take notes of something you're reading, listening to, watching, observing

Idea Notes: Review fleeting/lit. notes and write *one idea per note*

File Ideas: Find a place to save the new ideas amongst existing ideas, so you can find them again in the future

Reuse Ideas: Let ideas you've collected over time inspire your next essay

NIKLAS' GOAL: To facilitate ongoing conversations with himself over time.



REMEMBER EVERYTHING, ACHIEVE ANYTHING

Building a Second Brain

'Forte's ideas really work' SETH GODIN

'Well-written, cogent and useful' DAVID ALLEN

A Proven Method to Organise Your Digital Life and Unlock Your Creative Potential

TIAGO FORTE

'Completely reshaped how I think about information and how and why I take notes' DAN PINK

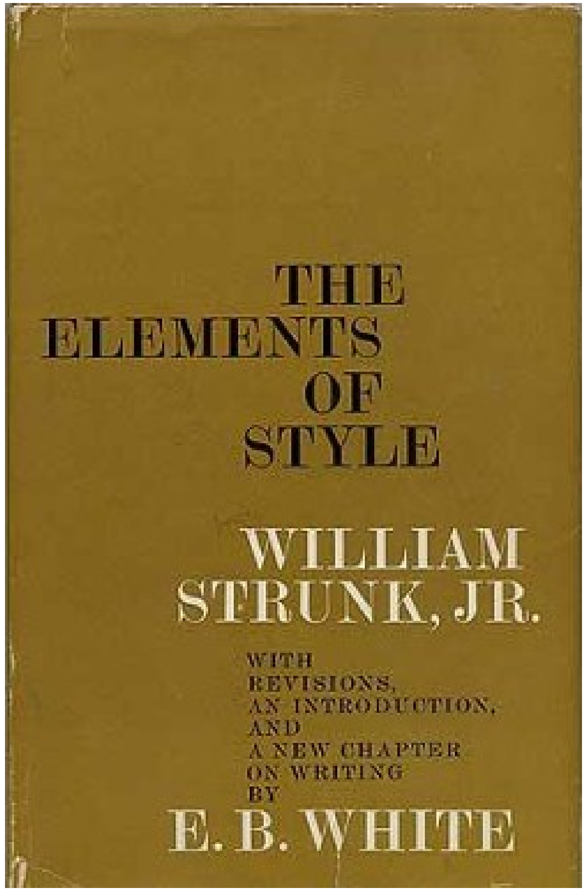
https://feeei.substack.com/p/the-dirty-lil-secret-about-my-note

Actual Text

- Papers are written in a particular *scientific style*
- Best advice is to read other papers
- And be aware of the grammar used
  - Active voice (“We ran the experiments”) versus passive voice (“The experiments were run”)
  - Tense
    - (“This paper presents” throughout, but “This paper has presented” in conclusion)
    - Past tense for experiments (“We ran...”)
  - Word choice (Latin “obtained” is better than Germanic “got”)
- Above all, clear and concise!

Elements of Style

- Useful to read to think about the elements of style
- Don’t follow every rule
- Most important: “omit needless words"
  - Shorter writing is usually better
  - Ex. remove “In order to”



THE
ELEMENTS
OF
STYLE

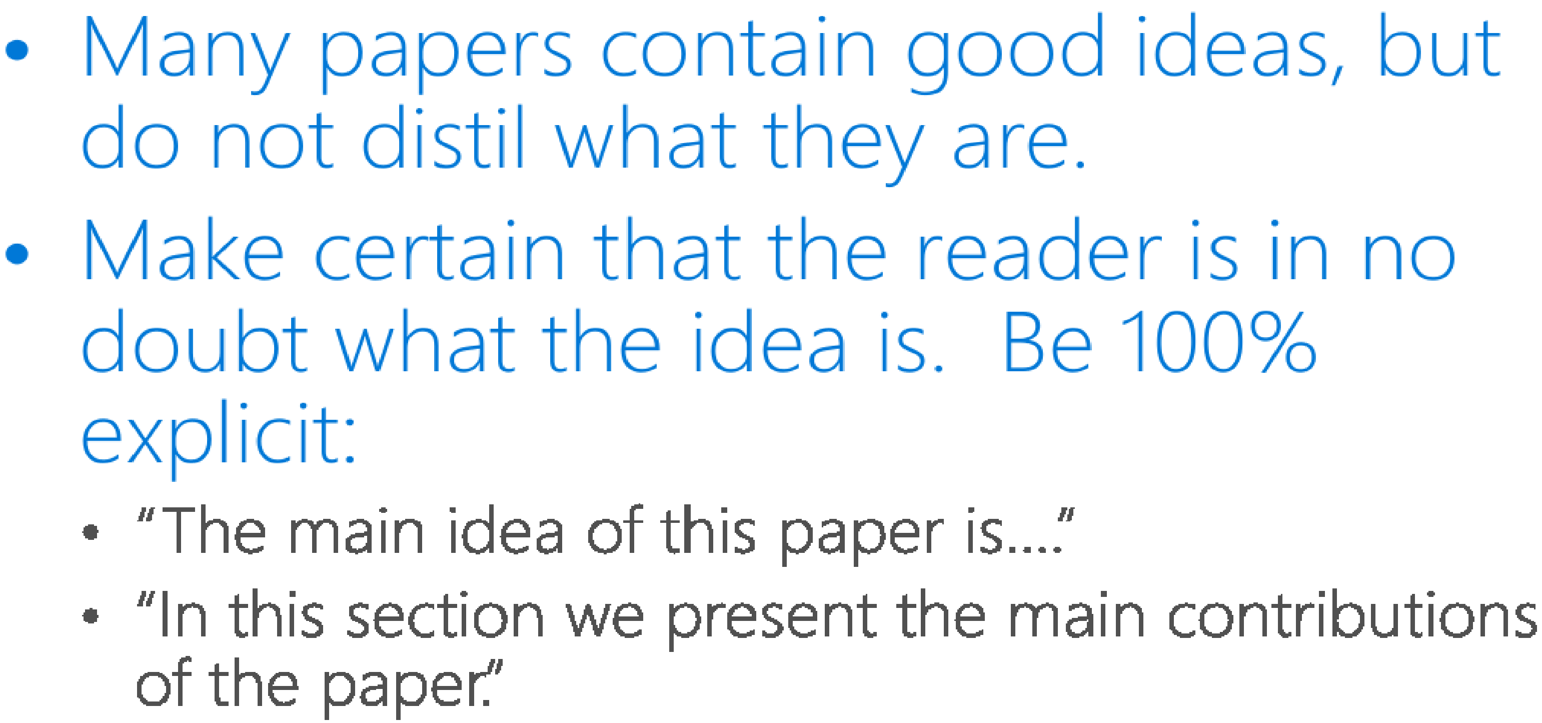
WILLIAM
STRUNK, JR.

WITH
REVISIONS,
AN INTRODUCTION,
AND
A NEW CHAPTER
ON WRITING
BY
E. B. WHITE

Plan Your Research

Plan Your Research

- Your paper should have just one “ping”: one **clear, sharp idea**



- Many papers contain good ideas, but do not distil what they are.
- Make certain that the reader is in no doubt what the idea is. Be 100% explicit:
  - “The main idea of this paper is....”
  - “In this section we present the main contributions of the paper.”

Finding Ideas

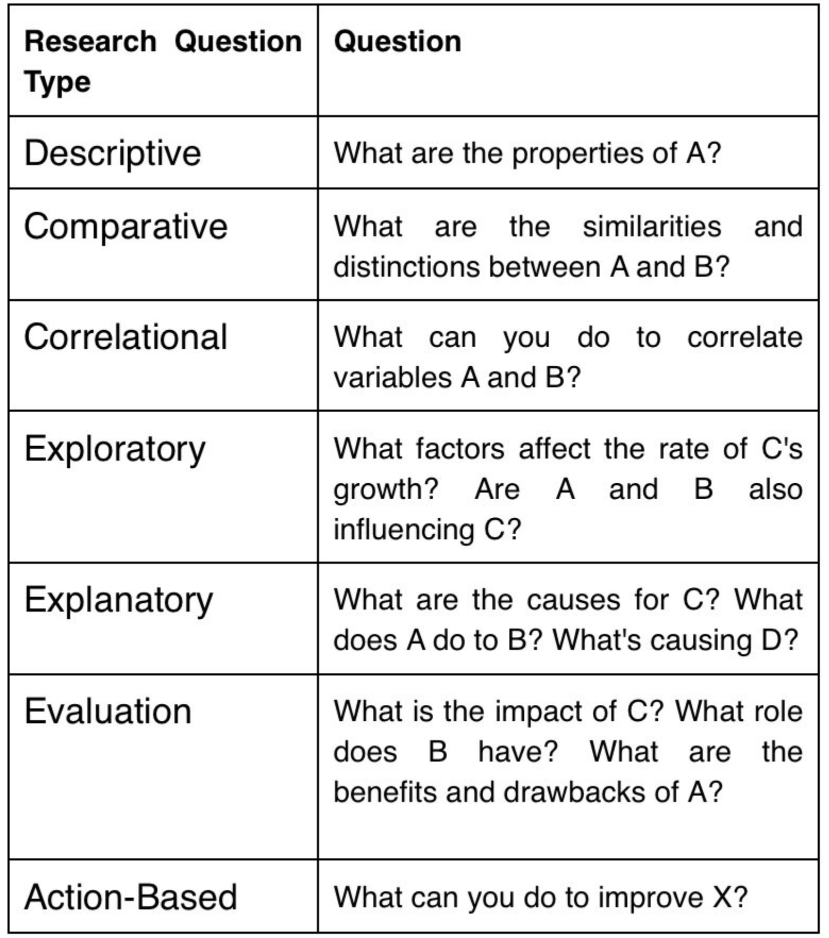
How do you come up with ideas?

- Read, review, discussions, conferences…
- You will have many (too many) ideas
- Real problem is how to:
  - Turn an idea into a research question (RQ)
  - Evaluate and validate that RQ

Research Questions

- RQs should be motivated by a concrete problem in your field.

- RQs can investigate:
  - what people are doing (repo mining, lit review)
  - what people want to do (interviews)
  - better approaches to do something (new techniques)



Research Question Type	Question
Descriptive	What are the properties of A?
Comparative	What are the similarities and distinctions between A and B?
Correlational	What can you do to correlate variables A and B?
Exploratory	What factors affect the rate of C's growth? Are A and B also influencing C?
Explanatory	What are the causes for C? What does A do to B? What's causing D?
Evaluation	What is the impact of C? What role does B have? What are the benefits and drawbacks of A?
Action-Based	What can you do to improve X?

Finding Research Question(s)

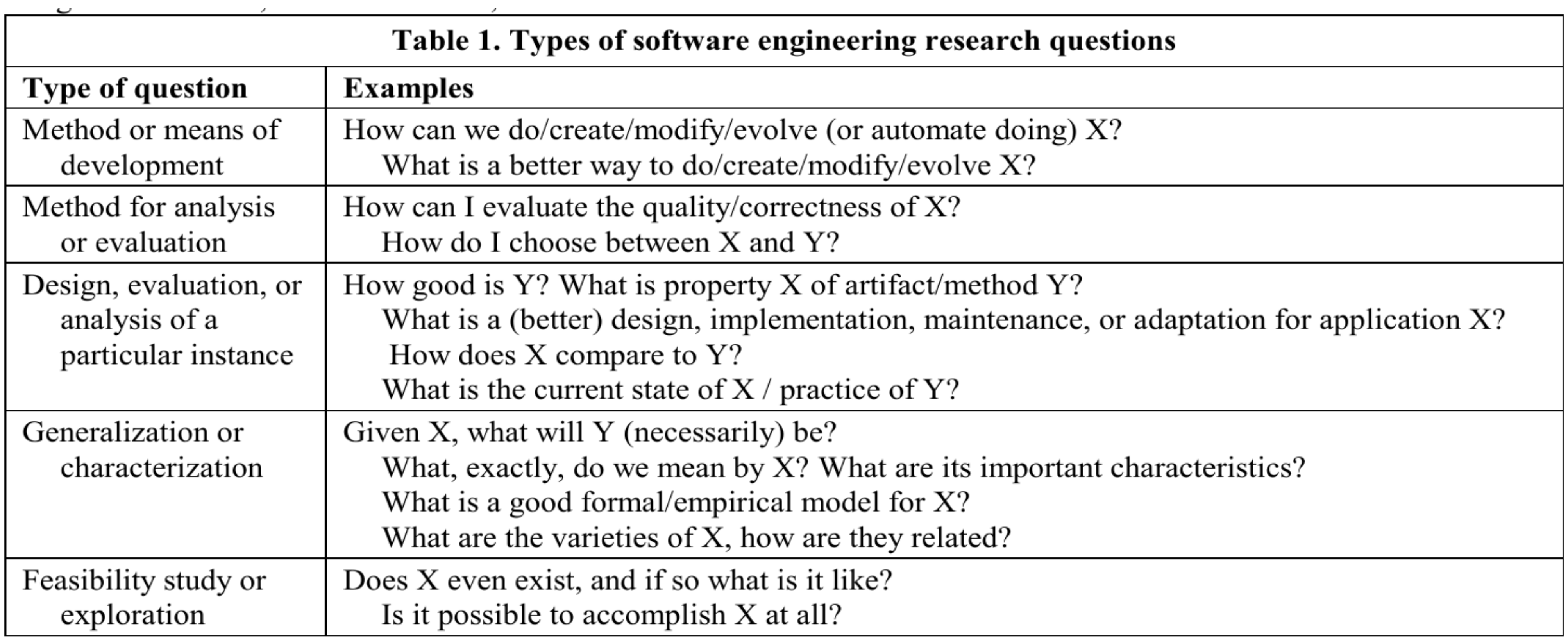
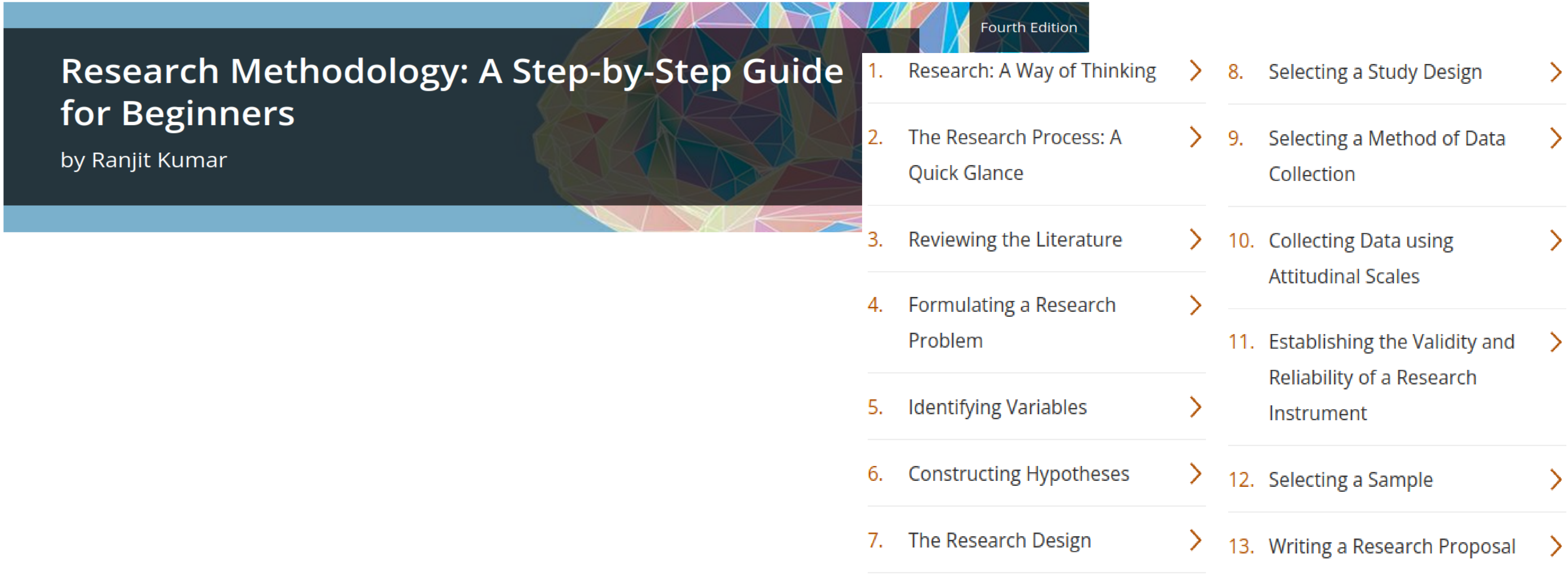


Table 1. Types of software engineering research questions

Type of question	**Examples**
Method or means of development	How can we do/create/modify/evolve (or automate doing) X? What is a better way to do/create/modify/evolve X?
Method for analysis or evaluation	How can I evaluate the quality/correctness of X? How do I choose between X and Y?
Design, evaluation, or analysis of a particular instance	How good is Y? What is property X of artifact/method Y? What is a (better) design, implementation, maintenance, or adaptation for application X? How does X compare to Y? What is the current state of X / practice of Y?
Generalization or characterization	Given X, what will Y (necessarily) be? What, exactly, do we mean by X? What are its important characteristics? What is a good formal/empirical model for X? What are the varieties of X, how are they related?
Feasibility study or exploration	Does X even exist, and if so what is it like? Is it possible to accomplish X at all?

Shaw, M. (2003, May). Writing good software engineering research papers. In 25th International Conference on Software Engineering, 2003. Proceedings. (pp. 726-736). IEEE.

Following Research Methodology



Fourth Edition

Research Methodology: A Step-by-Step Guide for Beginners

by Ranjit Kumar

Evaluating a Research Question

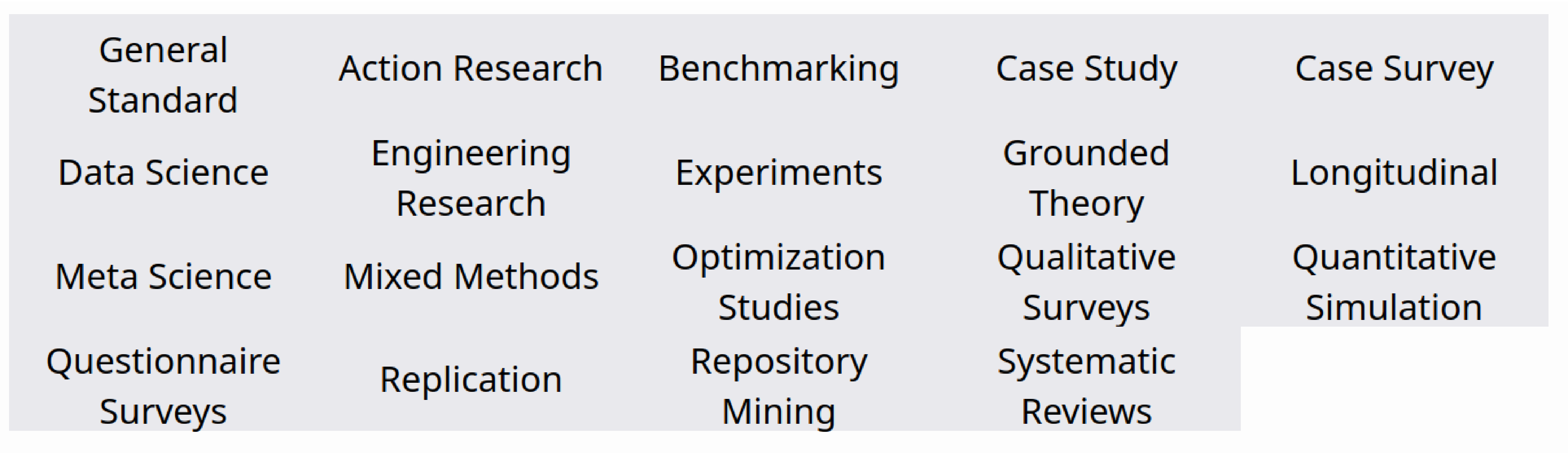
- Once we have RQ(s), the paper needs to clearly evaluate it.

- "Approach X is good at solving this problem"
  - How do you know that? Better than Y?

- "Users will like approach X"
  - How do you know that? Did you ask them?

- What are the metrics to evaluate your RQ?
  - Survey scores, precision/recall, seconds taken, confusion matrices, bar charts

Connecting RQ to Evaluation



General Standard	Action Research	Benchmarking	Case Study	Case Survey
Data Science	Engineering Research	Experiments	Grounded Theory	Longitudinal
Meta Science	Mixed Methods	Optimization Studies	Qualitative Surveys	Quantitative Simulation
Questionnaire Surveys	Replication	Repository Mining	Systematic Reviews	

ACM SIGSOFT Empirical Standards for Software Engineering
https://www2.sigsoft.org/EmpiricalStandards/docs/

Writing Proposals

Writing Proposals

- Proposals are different to papers
- Papers are about **explaining** an idea and your approach
- Proposals are about ***selling*** an idea and your approach

In a proposal, you are *asking for money*
You must have:

- A clear idea, which is important to fund
- Know what others have done
- Clear objectives
- Clear approach and a feasible timeline
- Written this in a short amount of text

Writing Proposals

1) Here is a **problem**
2) It's an **important problem** (evidence...)
3) We have a **promising idea** (evidence...)
4) We are a **world-class team** (evidence...)
5) Here is what we **hope to achieve**, and how we'll know if we have **succeeded**.
6) Here is a **plan** of how we're going to get from our idea to that destination
7) Give us the **money**. Please.

https://www.microsoft.com/en-us/research/academic-program/how-to-write-a-great-research-proposal/

Writing Proposals

1. Set the context for research or the problem to be solved (say that the field is important),
2. Say what you are going to do and why it is important.
3. Say how you are going to do it.
4. Talk about impact and future work.

https://www.uregina.ca/gradstudies/assets/docs/ppt/GraduateStudentConference/NSERC.pdf

NSERC Proposal Structure

Overview: Provide a detailed description of your proposed research project for the period during which you will hold the award. Be as specific as possible.

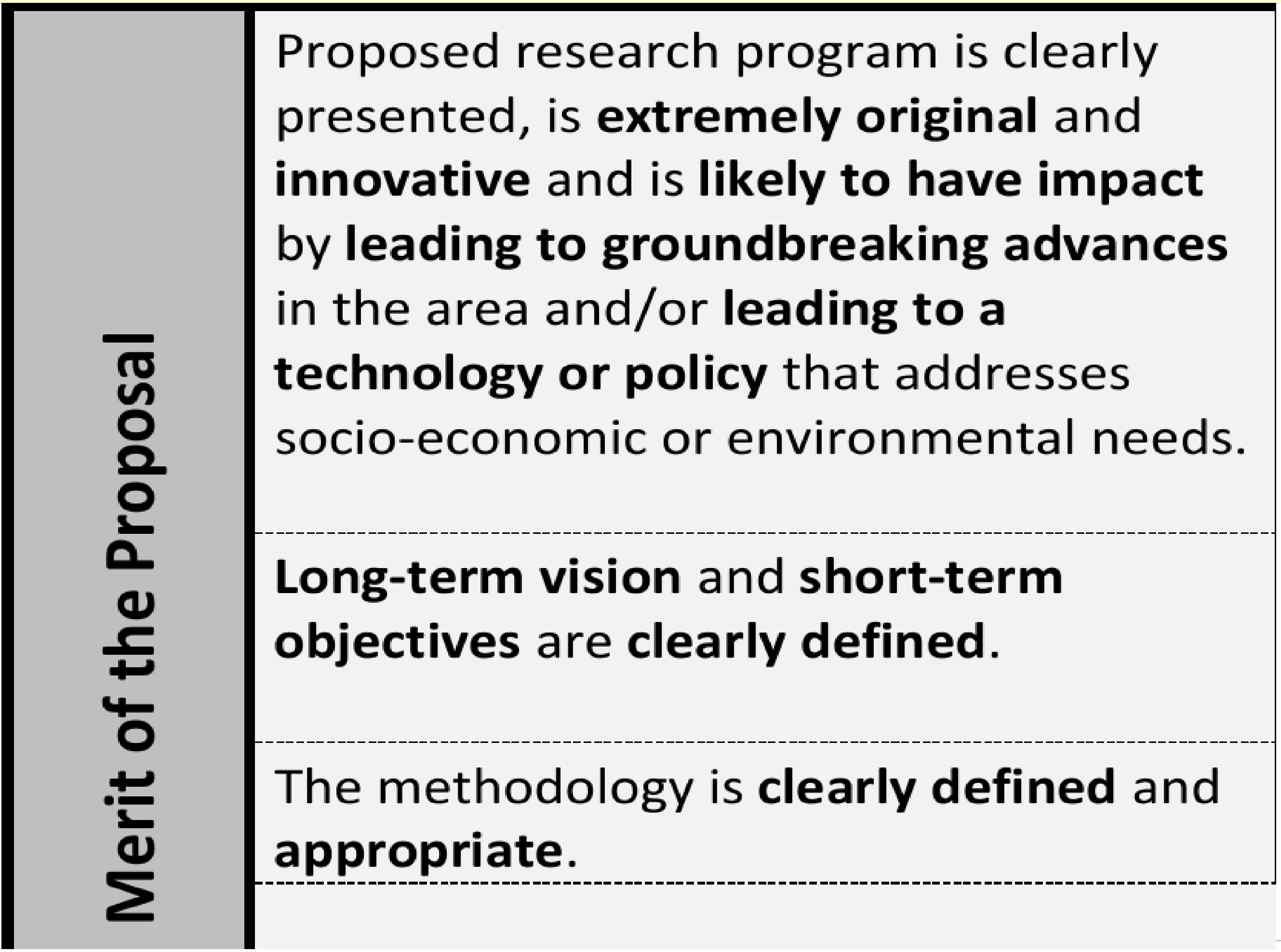
Background: Provide background information to position your proposed research within the context of current knowledge in the field.

Objectives and Hypothesis:
State the objectives and hypothesis,

Approach: and outline the experimental or theoretical approach to be taken (citing literature pertinent to the proposal) and the methods and procedures to be used.

Impact: State the significance of the proposed research to a field or fields in the health sciences, natural sciences and/or engineering or social sciences and/or humanities, as appropriate.

NSERC Discovery Evaluation



Merit of the Proposal	Proposed research program is clearly presented, is **extremely original** and **innovative** and is **likely to have impact** by **leading to groundbreaking advances** in the area and/or **leading to a technology or policy** that addresses socio-economic or environmental needs.
	Long-term vision and **short-term objectives** are **clearly defined**.
	The methodology is **clearly defined** and **appropriate**.

Proposal Examples

- https://research.viu.ca/cgs-m-examples
  https://figshare.com/articles/online_resource/Bernhardt-CGS-M-NSERC_pdf/17058125/1

Paper Structure

Paper Structure

- Depends on type of research
  - Literature survey, vision paper, proposed approach, etc.
- Depends on type of paper
  - Journals have ~20-30 pages, conferences have ~10 pages

1) Title
2) Abstract
3) Introduction
4) Background (for journal)
5) Approach
6) Evaluation set-up / results
7) Discussion
8) Related Work
9) Conclusion

Most Important Part?

1) Title
2) Abstract
3) Introduction
4) Background (for journal)
5) Approach
6) Evaluation set-up / results
7) Discussion
8) Related Work
9) Conclusion

Most Important Part for Me: Intro

1) Title
2) Abstract
3) Introduction
4) Background (for journal)
5) Approach
6) Evaluation set-up / results
7) Discussion
8) Related Work
9) Conclusion

Introduction:

a) Context to the problem

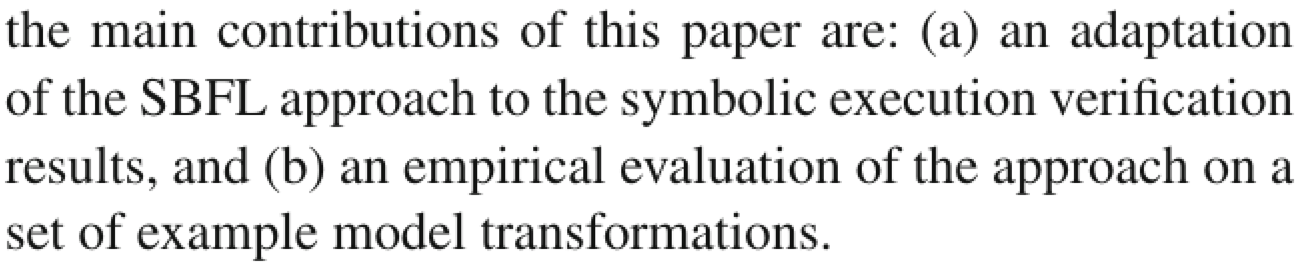
b) What’s the **specific problem** addressed**?**

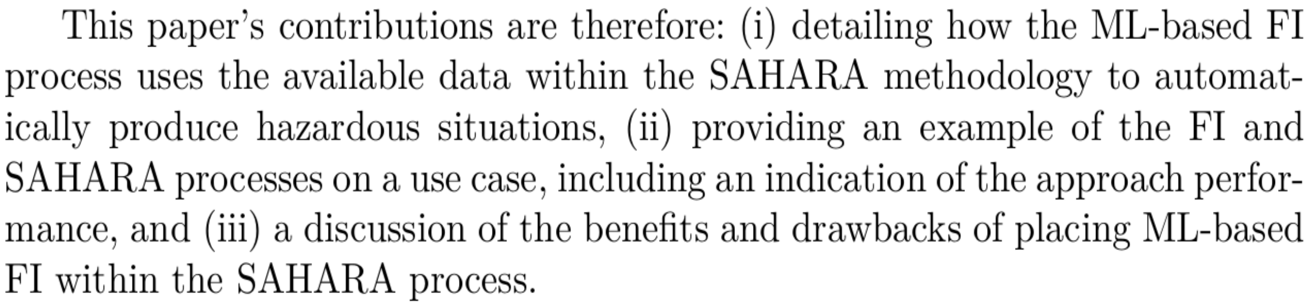
c) Brief summary of the **approach**

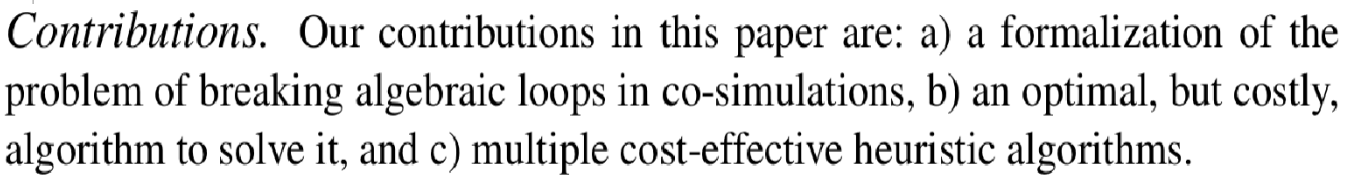
d) **<u>Contributions</u>** and **research questions**

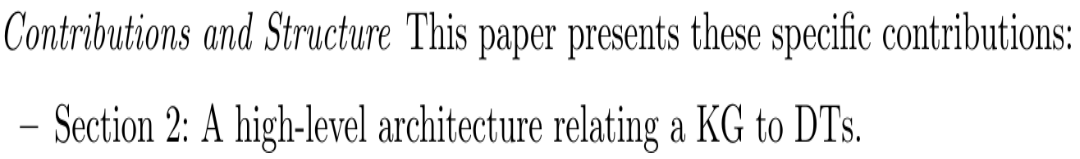
e) Structure of the paper

Examples of Contributions

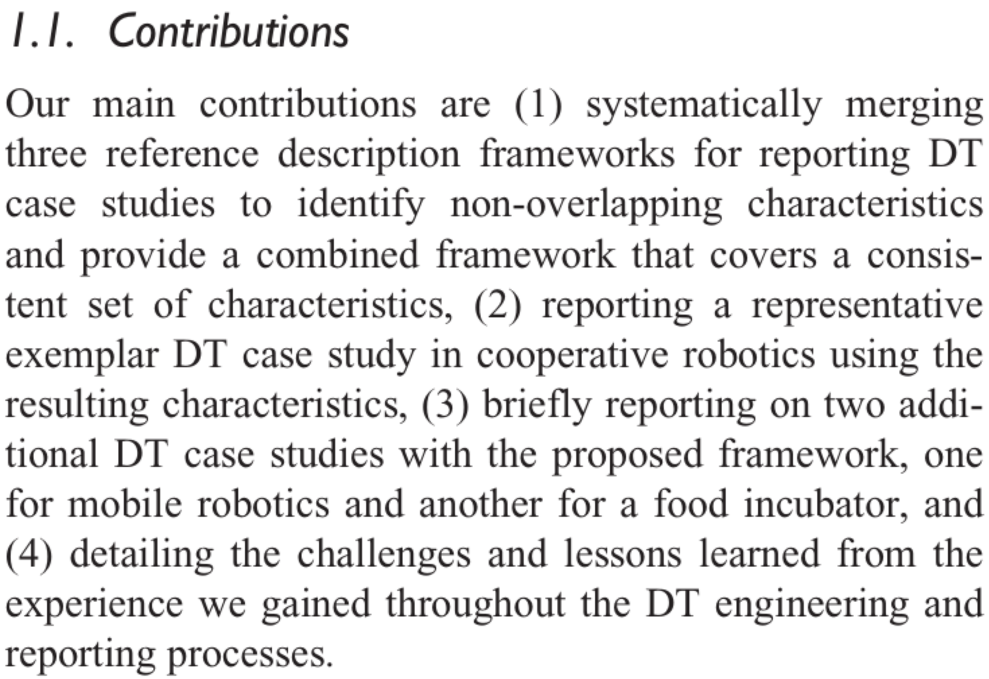

the main contributions of this paper are: (a) an adaptation of the SBFL approach to the symbolic execution verification results, and (b) an empirical evaluation of the approach on a set of example model transformations.


This paper's contributions are therefore: (i) detailing how the ML-based FI process uses the available data within the SAHARA methodology to automatically produce hazardous situations, (ii) providing an example of the FI and SAHARA processes on a use case, including an indication of the approach performance, and (iii) a discussion of the benefits and drawbacks of placing ML-based FI within the SAHARA process.


Contributions. Our contributions in this paper are: a) a formalization of the problem of breaking algebraic loops in co-simulations, b) an optimal, but costly, algorithm to solve it, and c) multiple cost-effective heuristic algorithms.

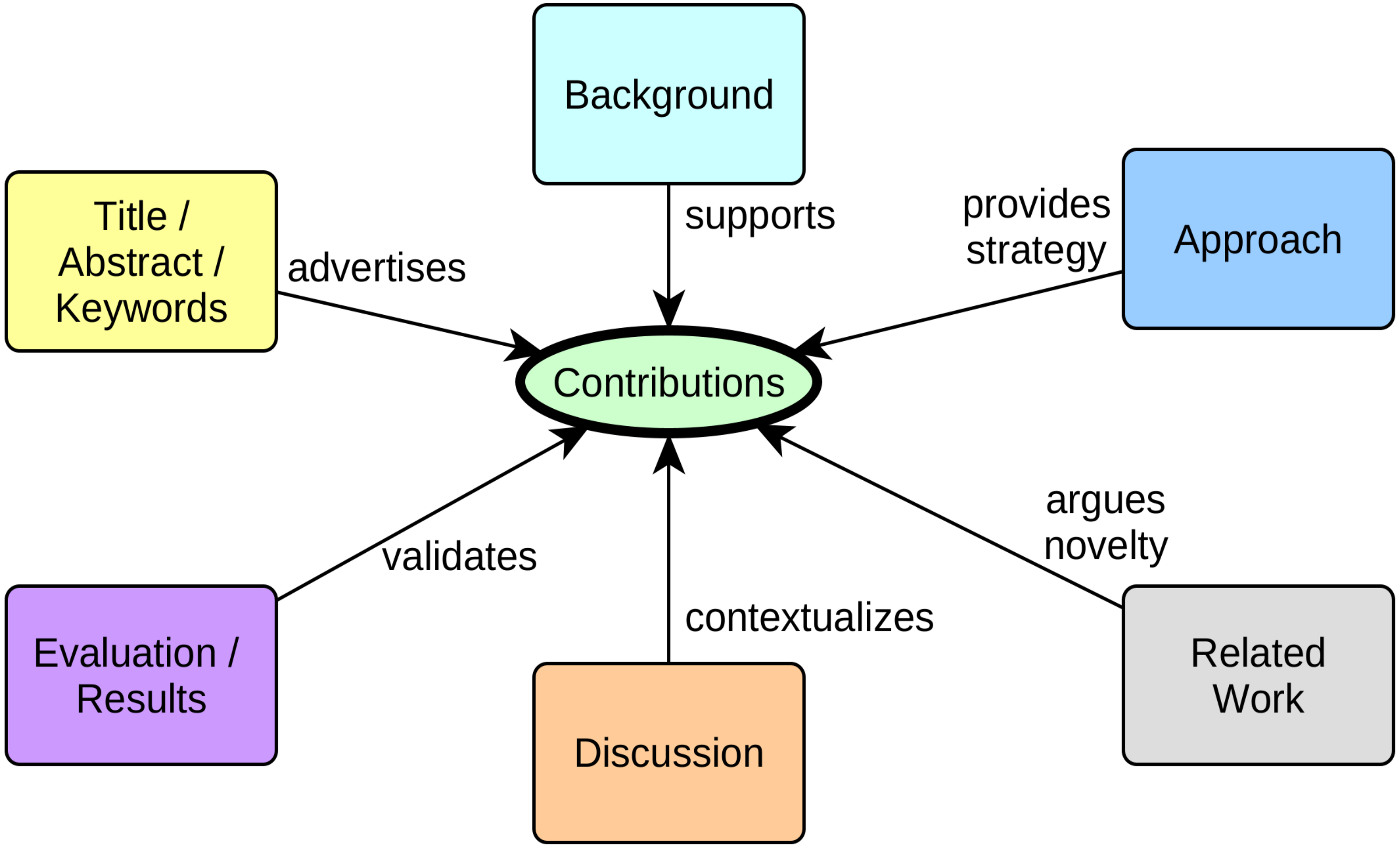

Contributions and Structure This paper presents these specific contributions:

– Section 2: A high-level architecture relating a KG to DTs.


1.1. Contributions

Our main contributions are (1) systematically merging three reference description frameworks for reporting DT case studies to identify non-overlapping characteristics and provide a combined framework that covers a consistent set of characteristics, (2) reporting a representative exemplar DT case study in cooperative robotics using the resulting characteristics, (3) briefly reporting on two additional DT case studies with the proposed framework, one for mobile robotics and another for a food incubator, and (4) detailing the challenges and lessons learned from the experience we gained throughout the DT engineering and reporting processes.

Paper Centres Around Contributions

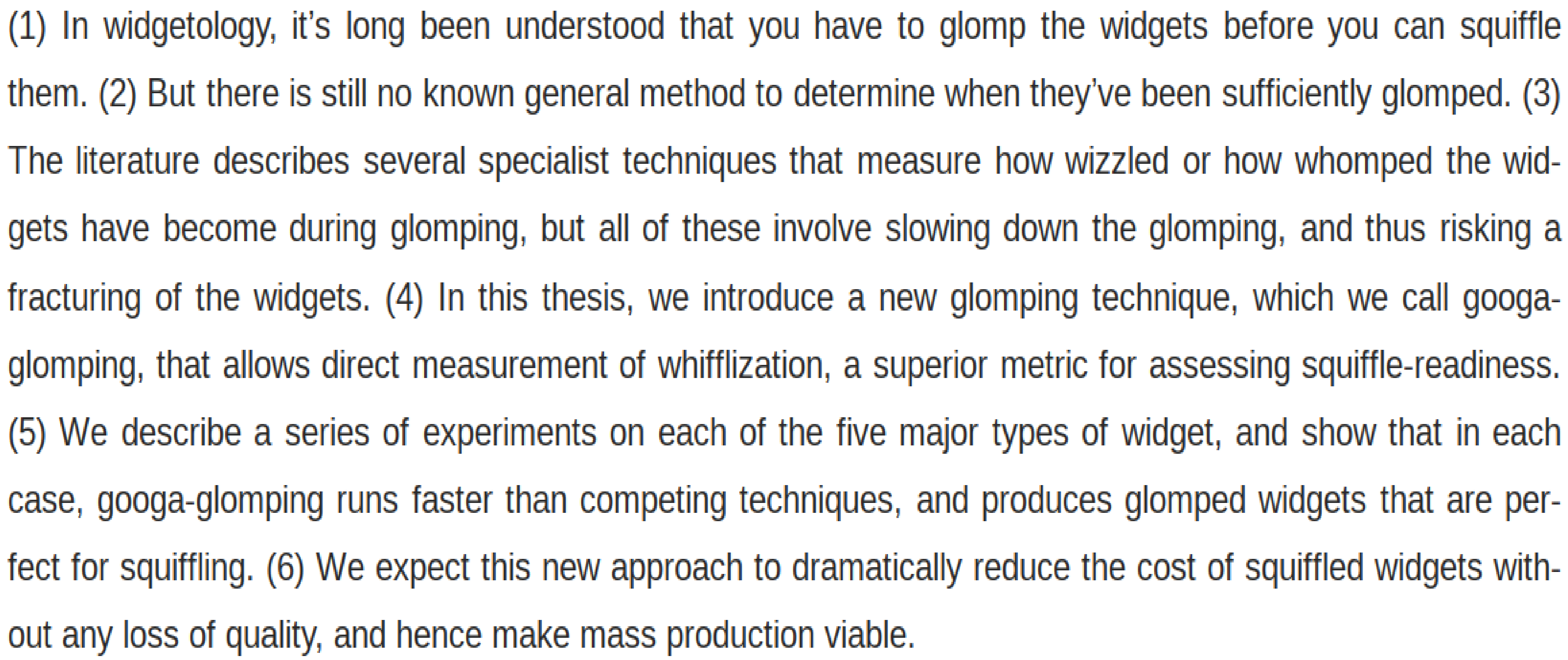


How to write a scientific abstract

1) In one sentence, what's the topic?
2) State the problem you tackle
3) Summarize (in one sentence) why nobody else has adequately answered the research question yet.
4) Explain, in one sentence, how you tackled the research question.
5) In one sentence, how did you go about doing the research that follows from your big idea.
6) As a single sentence, what's the key impact of your research?

https://www.easterbrook.ca/steve/2010/01/how-to-write-a-scientific-abstract-in-six-easy-steps/

How to write a scientific abstract

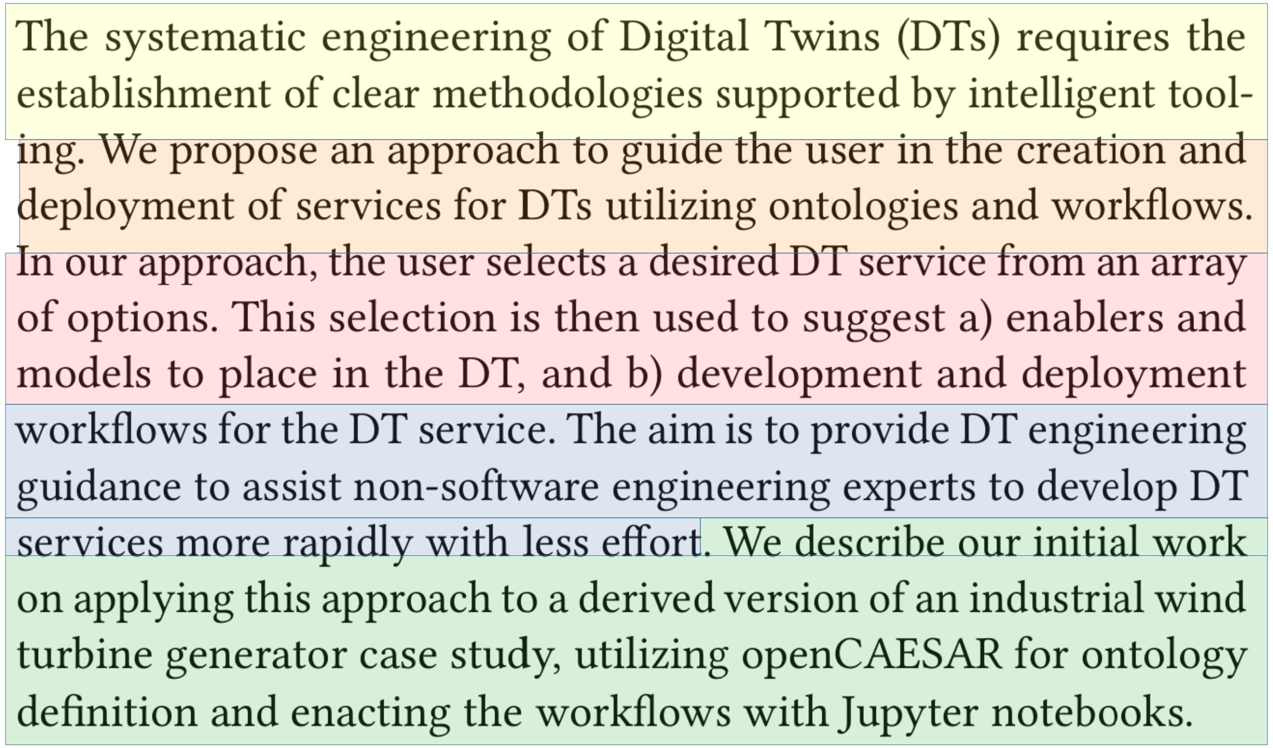


(1) In widgetology, it’s long been understood that you have to glomp the widgets before you can squiffle them. (2) But there is still no known general method to determine when they’ve been sufficiently glomped. (3) The literature describes several specialist techniques that measure how wizzled or how whomped the widgets have become during glomping, but all of these involve slowing down the glomping, and thus risking a fracturing of the widgets. (4) In this thesis, we introduce a new glomping technique, which we call googa-glomping, that allows direct measurement of whifflization, a superior metric for assessing squiffle-readiness. (5) We describe a series of experiments on each of the five major types of widget, and show that in each case, googa-glomping runs faster than competing techniques, and produces glomped widgets that are perfect for squiffling. (6) We expect this new approach to dramatically reduce the cost of squiffled widgets without any loss of quality, and hence make mass production viable.

https://www.easterbrook.ca/steve/2010/01/how-to-write-a-scientific-abstract-in-six-easy-steps/

Title / Abstract / Keywords

Towards Ontological Service-Driven Engineering of Digital Twins



The systematic engineering of Digital Twins (DTs) requires the establishment of clear methodologies supported by intelligent tooling. We propose an approach to guide the user in the creation and deployment of services for DTs utilizing ontologies and workflows. In our approach, the user selects a desired DT service from an array of options. This selection is then used to suggest a) enablers and models to place in the DT, and b) development and deployment workflows for the DT service. The aim is to provide DT engineering guidance to assist non-software engineering experts to develop DT services more rapidly with less effort. We describe our initial work on applying this approach to a derived version of an industrial wind turbine generator case study, utilizing openCAESAR for ontology definition and enacting the workflows with Jupyter notebooks.

Keywords:
digital twins, ontologies, DT services, wind turbine testing, guided software engineering, recommendation, workflows

Oakes et al (2024). Towards Ontological Service-Driven Engineering of Digital Twins.
In 2023 Int. Conf. on Engineering Digital Twins (EDTConf). IEEE.

Intro

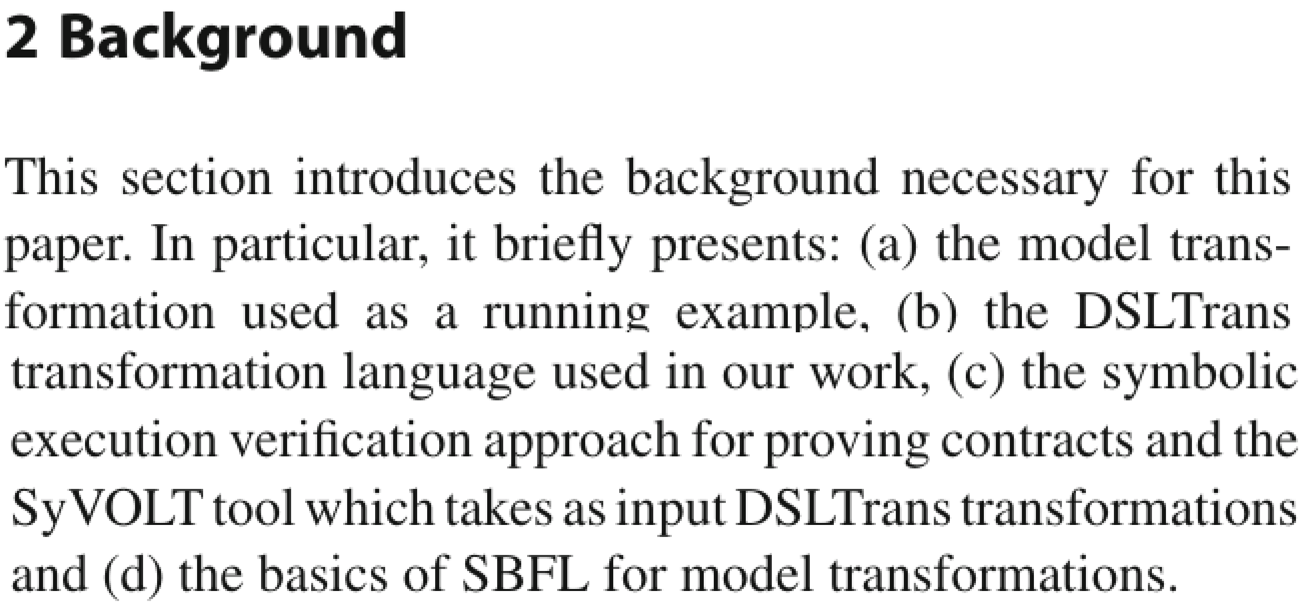
Introduction:
a) Context to the problem
b) What's the **specific problem** addressed**?**
c) Brief summary of the **approach**
d) **<u>Contributions</u>** and **research questions**
e) Structure of the paper

Tips:

- Be explicit and specific with these parts
- Start the paper with a draft of the intro
- Read other papers and note their intro style

Background

- Usually only for journal articles
- Mention the fundamental topics in the article needed to understand the approach



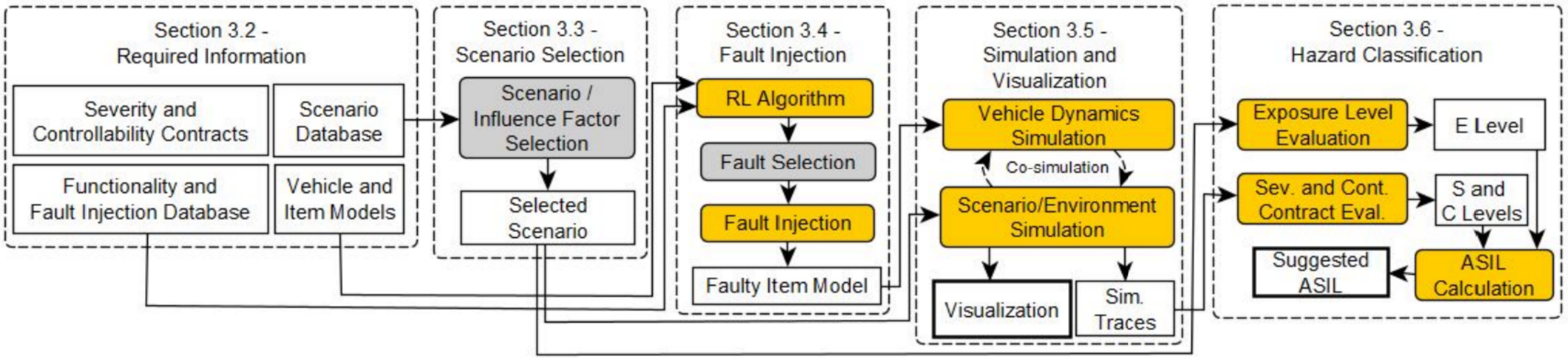
2 Background

This section introduces the background necessary for this paper. In particular, it briefly presents: (a) the model transformation used as a running example, (b) the DSLTrans transformation language used in our work, (c) the symbolic execution verification approach for proving contracts and the SyVOLT tool which takes as input DSLTrans transformations and (d) the basics of SBFL for model transformations.

Oakes et al (2023). Fault localization in DSLTrans model transformations by combining symbolic execution and spectrum-based analysis. Software and Systems Modeling, 1-27.

Approach

- Detail the proposed approach, and try to include an overview diagram

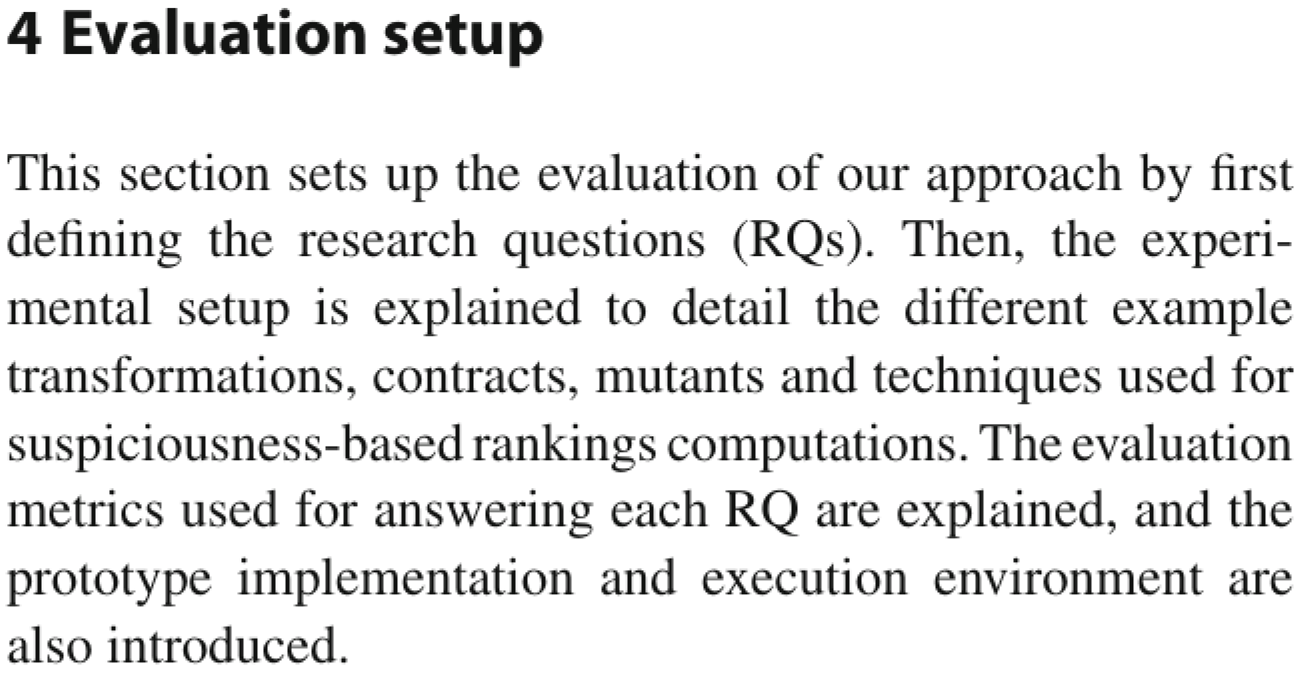


Section 3.2 - Required Information
Severity and Controllability Contracts
Scenario Database
Functionality and Fault Injection Database
Vehicle and Item Models
Section 3.3 - Scenario Selection
Scenario / Influence Factor Selection
Selected Scenario
Section 3.4 - Fault Injection
RL Algorithm
Fault Selection
Fault Injection
Faulty Item Model
Section 3.5 - Simulation and Visualization
Vehicle Dynamics Simulation
Co-simulation
Scenario/Environment Simulation
Visualization
Sim. Traces
Section 3.6 - Hazard Classification
Exposure Level Evaluation
E Level
Sev. and Cont. Contract Eval.
S and C Levels
Suggested ASIL
ASIL Calculation

Fig. 3: The overall SAHARA architecture and workflow (adapted from [21]). Yellow blocks are automatic actions, and gray blocks are manual actions.

Oakes et al. (2021, August). Machine Learning-Based Fault Injection for Hazard Analysis and Risk Assessment. In International Conference on Computer Safety, Reliability, and Security (pp. 178-192).

Evaluation

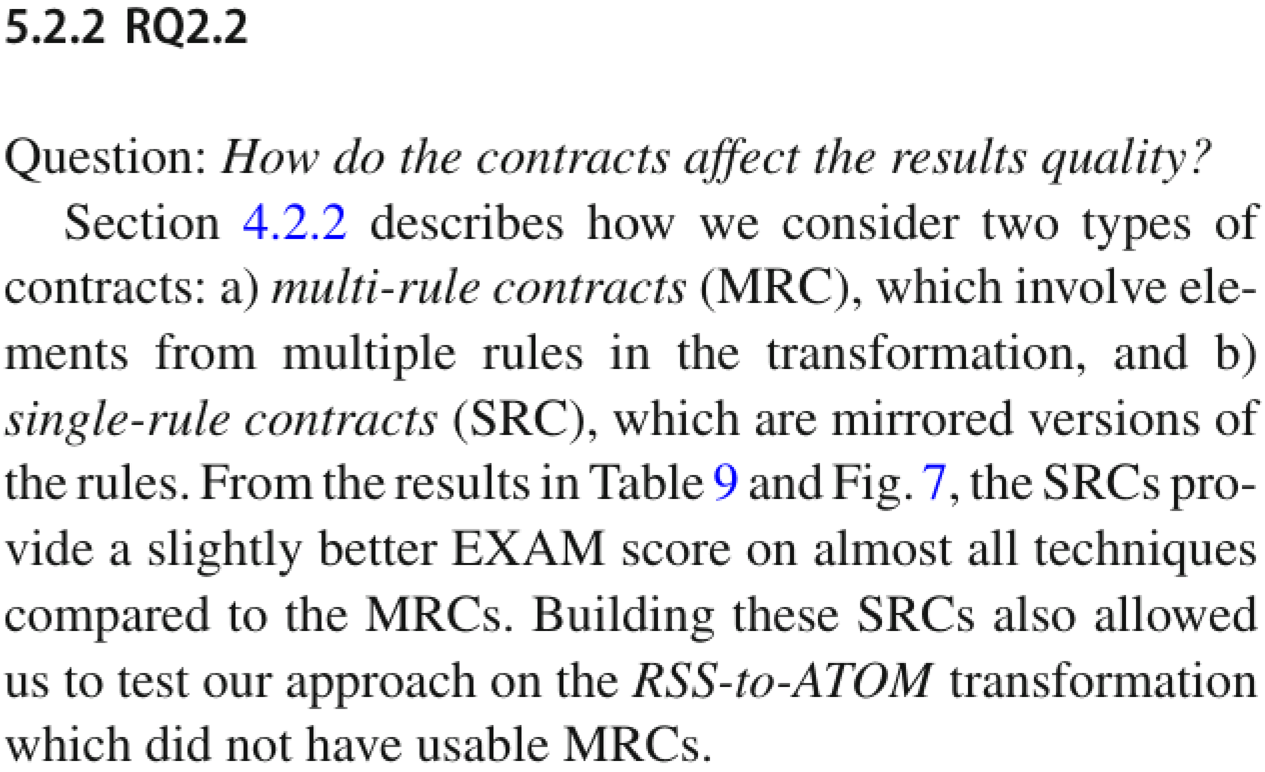


4 Evaluation setup

This section sets up the evaluation of our approach by first defining the research questions (RQs). Then, the experimental setup is explained to detail the different example transformations, contracts, mutants and techniques used for suspiciousness-based rankings computations. The evaluation metrics used for answering each RQ are explained, and the prototype implementation and execution environment are also introduced.

Oakes et al (2023). Fault localization in DSLTrans model transformations by combining symbolic execution and spectrum-based analysis. Software and Systems Modeling, 1-27.

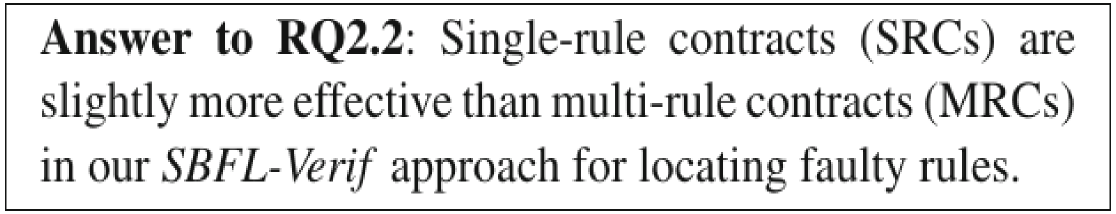
Results



5.2.2 RQ2.2

Question: *How do the contracts affect the results quality?*

Section 4.2.2 describes how we consider two types of contracts: a) *multi-rule contracts* (MRC), which involve elements from multiple rules in the transformation, and b) *single-rule contracts* (SRC), which are mirrored versions of the rules. From the results in Table 9 and Fig. 7, the SRCs provide a slightly better EXAM score on almost all techniques compared to the MRCs. Building these SRCs also allowed us to test our approach on the *RSS-to-ATOM* transformation which did not have usable MRCs.



Answer to RQ2.2: Single-rule contracts (SRCs) are slightly more effective than multi-rule contracts (MRCs) in our *SBFL-Verif* approach for locating faulty rules.

Oakes et al (2023). Fault localization in DSLTrans model transformations by combining symbolic execution and spectrum-based analysis. Software and Systems Modeling, 1-27.

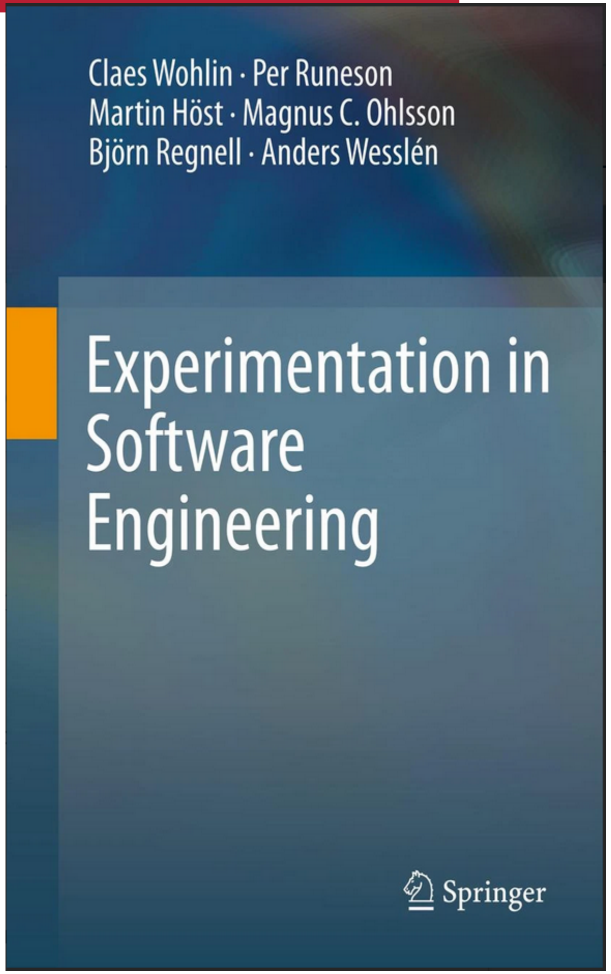
Discussion

- Often missing from papers and could lead to rejection

- My structure:
  - **Benefits** of the approach
    - Based on the evaluation, argue that results are good
  - **Limitations** of the approach
    - Anticipate reviewer's objections
  - **Threats to validity**

Threats to Validity

- **Conclusion validity**
  - Threats to obtaining conclusions from results
  - Ex. Are stats reliable and powerful enough, any random heterogeneity of subjects, repeatable?
- **Internal validity**
  - Threats that might affect the results
  - Ex. Sufficient metrics to evaluate approach, ensuring that approach was performed correctly
- **Construct validity**
  - Threats of the approach not matching the theory
  - Ex. Metrics not matching problem to be studied
- **External validity**
  - Threats to generalization, Ex. Subset of population studied



Claes Wohlin · Per Runeson
Martin Höst · Magnus C. Ohlsson
Björn Regnell · Anders Wesslén

Experimentation in Software Engineering

Springer

Wohlin, C., Runeson, P., Höst, M., Ohlsson, M.C., Regnell, B.: Experimentation in Software Engineering. Springer, Berlin (2012)

Related Work

Purpose: Shows the reader that the contributions are novel

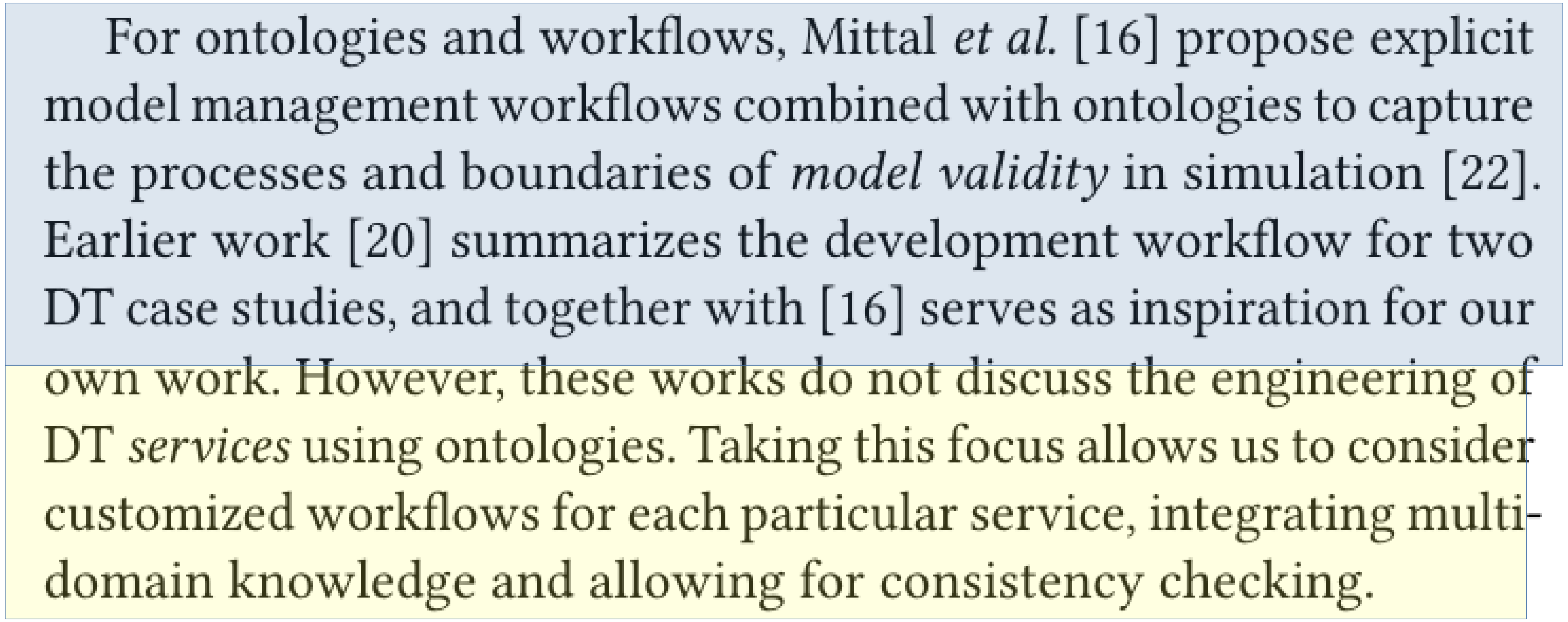
Therefore, you must:

a) summarize the work of others, and

b) **explain how your work is different/better**

Put related work at the end of the paper
Reader understands your approach/results, and you can be more specific

Related Work Example

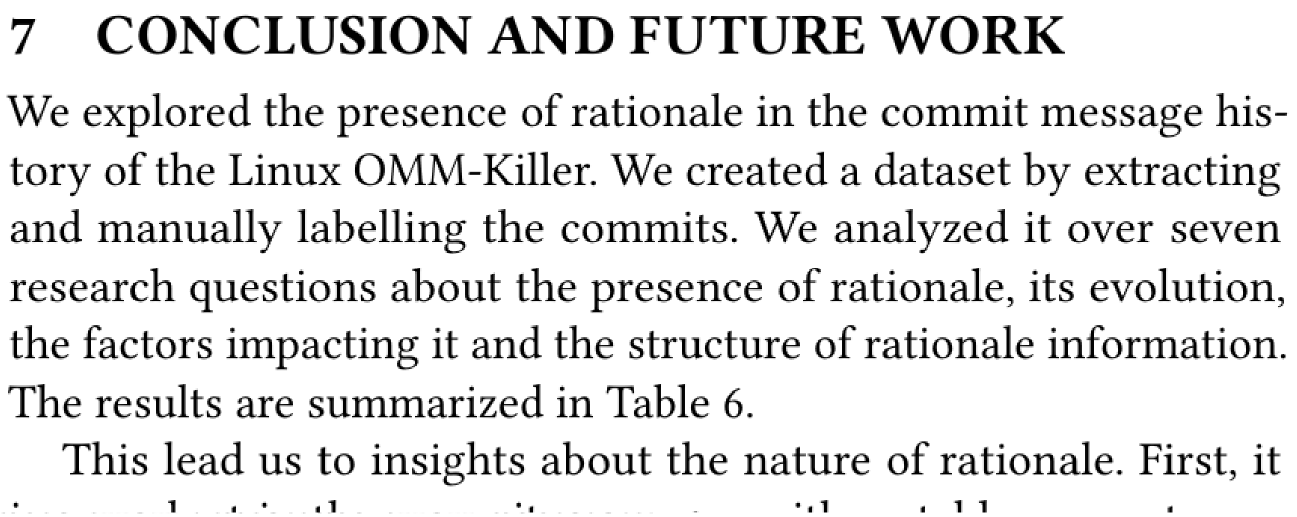


For ontologies and workflows, Mittal *et al.* [16] propose explicit model management workflows combined with ontologies to capture the processes and boundaries of *model validity* in simulation [22]. Earlier work [20] summarizes the development workflow for two DT case studies, and together with [16] serves as inspiration for our own work. However, these works do not discuss the engineering of DT *services* using ontologies. Taking this focus allows us to consider customized workflows for each particular service, integrating multi-domain knowledge and allowing for consistency checking.

Oakes et al (2024). Towards Ontological Service-Driven Engineering of Digital Twins.
In 2023 Int. Conf. on Engineering Digital Twins (EDTConf). IEEE.

Conclusion

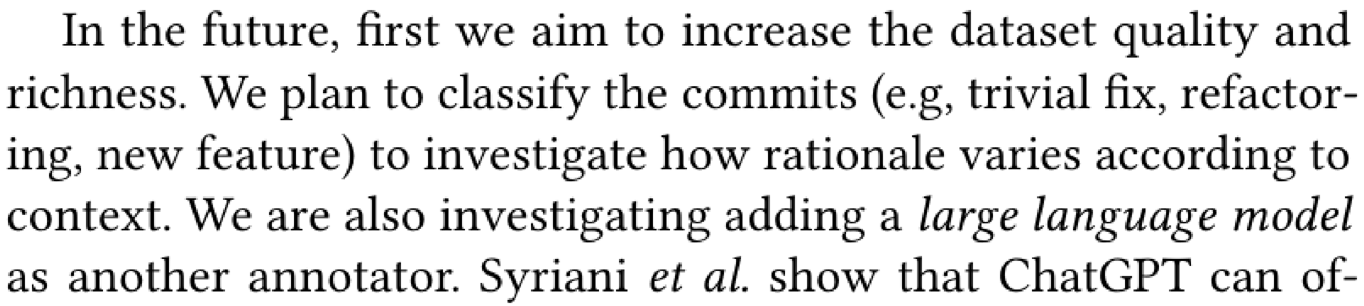
- Summarize the paper
  - Approach, insights
  - More detail than intro, as reader understands more
- Future work
  - Mention that limitations will be addressed, and how
  - You don’t have to actually do the future work



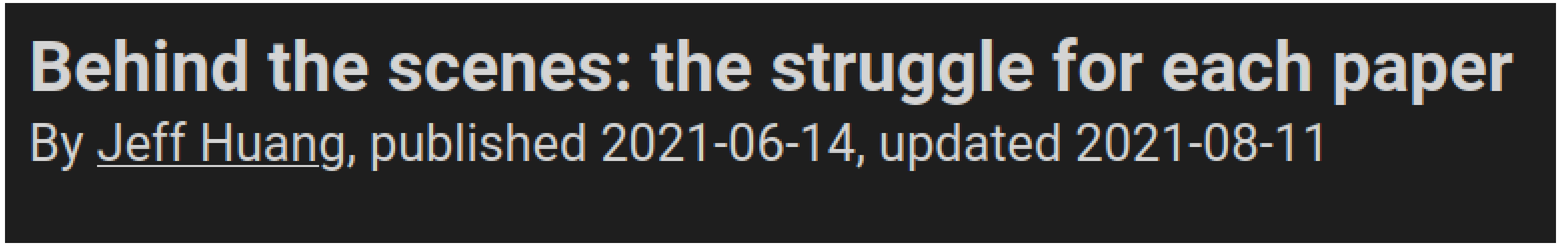
7 CONCLUSION AND FUTURE WORK

We explored the presence of rationale in the commit message history of the Linux OMM-Killer. We created a dataset by extracting and manually labelling the commits. We analyzed it over seven research questions about the presence of rationale, its evolution, the factors impacting it and the structure of rationale information. The results are summarized in Table 6.

This lead us to insights about the nature of rationale. First, it



In the future, first we aim to increase the dataset quality and richness. We plan to classify the commits (e.g, trivial fix, refactoring, new feature) to investigate how rationale varies according to context. We are also investigating adding a *large language model* as another annotator. Syriani *et al.* show that ChatGPT can of-



Behind the scenes: the struggle for each paper

By Jeff Huang, published 2021-06-14, updated 2021-08-11

https://jeffhuang.com/struggle_for_each_paper/

Takeaways:

- Papers are a mix of hard work, circumstances, and luck
- Papers depend on helping hands
- Good papers get rejected

Lab Session

In groups, discuss your paper

1) What are the contributions?
2) Is there an evaluation?
3) Do they highlight limitations?
4) What is the language like? Is it clear/concise?
5) What do you like about the writing?
6) What do you not like about the writing?

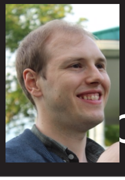
THANK YOU!

Topics:

- 1. Plan your writing
- 2. Plan your research
- 3. Paper structure
- 4. Proposal structure

POLYTECHNIQUE
MONTRÉAL

TECHNOLOGICAL
UNIVERSITY



Dr. Bentley James Oakes
bentleyjoakes.github.io