

# Putting Software Testing Terminology to the Test

## M.A.Sc. Seminar

Samuel Crawford, B.Eng.

McMaster University  
Department of Computing and Software

Fall 2024

# Table of Contents

## 1 Introduction

- The Need for Standardized Terminology
- The Lack of Standardized Terminology

## 2 Project

- Drasil
- Why Test Generated Code?
- Next Steps

## 3 References

# Table of Contents

## 1 Introduction

- The Need for Standardized Terminology
- The Lack of Standardized Terminology

## 2 Project

- Drasil
- Why Test Generated Code?
- Next Steps

## 3 References

# The Need for Standardized Terminology

- Engineering is applied science
- Scientific fields use precise terminology
- Therefore, the same should be true of software engineering!

# The Need for Standardized Terminology

- Engineering is applied science
- Scientific fields use precise terminology
- Therefore, the same should be true of software engineering!
- Imagine if other fields used unclear, inconsistent, and incorrect terminology:
  - Force
  - Isotope
  - Phalange

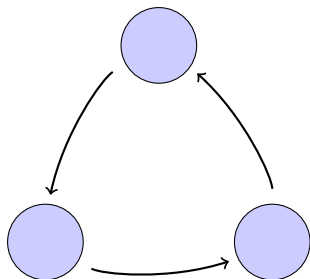
# The Need for Standardized Terminology

- Engineering is applied science
- Scientific fields use precise terminology
- Therefore, the same should be true of software engineering!
- Imagine if other fields used unclear, inconsistent, and incorrect terminology:
  - Force
  - Isotope
  - Phalange

If software engineering holds code to high standards of clarity, consistency, and robustness, the same should apply to its supporting literature!

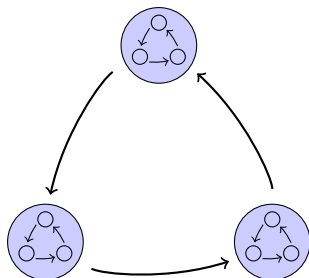
## Interorganizational

Schools, companies, etc.



## Interorganizational

Schools, companies, etc.



## Intraorganizational

Kaner et al. (2011, p. 7) say  
“complete testing” could require the  
tester to:

- discover “every bug”,
- exhaust the time allocated,
- implement every planned test,
- ...



# The Lack of Standardized Terminology

## “The Problem”

- Unfortunately, a search for a systematic, rigorous, and complete taxonomy for software testing revealed that the existing ones are inadequate:
  - Tebes et al. (2020) focus on *parts* of the testing process (e.g., test goal, testable entity),
  - Souza et al. (2017) prioritize organizing testing approaches over defining them, and
  - Unterkalmsteiner et al. (2014) focus on the “information linkage or transfer” (p. A:6) between requirements engineering and software testing.

# Unstandardized Standards

- The structure of tours is:
  - quite general (ISO/IEC and IEEE, 2022, p. 34)

- The structure of tours is:
  - quite general (ISO/IEC and IEEE, 2022, p. 34)
  - “organized around a special focus” (Hamburg and Mogyorodi, 2024)

- Load testing is performed with:
  - loads “between anticipated conditions of low, typical, and peak usage” (ISO/IEC and IEEE, 2022, p. 5)

- Load testing is performed with:
  - loads “between anticipated conditions of low, typical, and peak usage” (ISO/IEC and IEEE, 2022, p. 5)
  - loads that are as large as possible (Patton, 2006, p. 86)

- Alpha testing is performed by:
  - “users within the organization developing the software” (ISO/IEC and IEEE, 2017, p. 17)

- Alpha testing is performed by:
  - “users within the organization developing the software” (ISO/IEC and IEEE, 2017, p. 17)
  - “a small, selected group of potential users” (Washizaki, 2024, p. 5-8)

- Alpha testing is performed by:
  - “users within the organization developing the software” (ISO/IEC and IEEE, 2017, p. 17)
  - “a small, selected group of potential users” (Washizaki, 2024, p. 5-8)
  - “roles outside the development organization” conducted “in the developer’s test environment” (Hamburg and Mogyorodi, 2024)



# Unstandardized Standards

- Alpha testing is performed by:
  - “users within the organization developing the software” (ISO/IEC and IEEE, 2017, p. 17)
  - “a small, selected group of potential users” (Washizaki, 2024, p. 5-8)
  - “roles outside the development organization” conducted “in the developer’s test environment” (Hamburg and Mogyorodi, 2024)

“Okay testing team, we want to conduct alpha testing on our product. What’s our timeline? Budget? Sample size?”

# Table of Contents

## 1 Introduction

- The Need for Standardized Terminology
- The Lack of Standardized Terminology

## 2 Project

- Drasil
- Why Test Generated Code?
- Next Steps

## 3 References

# Problem Statement

- Currently, there is no way to verify Drasil's output

# Problem Statement

- Currently, there is no way to verify Drasil's output
- Drasil is "tested" by comparing generated artifacts to stable

# Problem Statement

- Currently, there is no way to verify Drasil's output
- Drasil is "tested" by comparing generated artifacts to stable
- This does not actually say anything about Drasil's output!

# Purpose Statement

- The purpose of this research is to implement test case generation to verify generated code
- These test cases will be generated from information within Drasil

# Purpose Statement

- The purpose of this research is to implement test case generation to verify generated code
- These test cases will be generated from information within Drasil
- Why use test cases for verification as opposed to, say, consistency/correctness checks?

# Purpose Statement

- The purpose of this research is to implement test case generation to verify generated code
- These test cases will be generated from information within Drasil
- Why use test cases for verification as opposed to, say, consistency/correctness checks?
  - 1 A more well-defined, Master's level scope



# Purpose Statement

- The purpose of this research is to implement test case generation to verify generated code
- These test cases will be generated from information within Drasil
- Why use test cases for verification as opposed to, say, consistency/correctness checks?
  - 1 A more well-defined, Master's level scope
  - 2 Targets a more complex artifact that is harder to verify

# Purpose Statement

- The purpose of this research is to implement test case generation to verify generated code
- These test cases will be generated from information within Drasil
- Why use test cases for verification as opposed to, say, consistency/correctness checks?
  - 1 A more well-defined, Master's level scope
  - 2 Targets a more complex artifact that is harder to verify
  - 3 Gives Drasil another "bragging point"!

# Why Test Generated Code?

If the code is being generated from a stable knowledge base, then it should be correct. Why waste effort testing it?

# Why Test Generated Code?

If the code is being generated from a stable knowledge base, then it should be correct. Why waste effort testing it?

- 1 The knowledge base is not actually "stable" yet

# Why Test Generated Code?

If the code is being generated from a stable knowledge base, then it should be correct. Why waste effort testing it?

- 1 The knowledge base is not actually "stable" yet
- 2 There are plenty of places for a mistake to be introduced

# Why Test Generated Code?

If the code is being generated from a stable knowledge base, then it should be correct. Why waste effort testing it?

- 1 The knowledge base is not actually "stable" yet
- 2 There are plenty of places for a mistake to be introduced
- 3 Testing provides a greater degree of confidence in Drasil's capabilities

# Why Test Generated Code?

If the code is being generated from a stable knowledge base, then it should be correct. Why waste effort testing it?

- 1 The knowledge base is not actually "stable" yet
- 2 There are plenty of places for a mistake to be introduced
- 3 Testing provides a greater degree of confidence in Drasil's capabilities
- 4 Generating code for testing allows for it to be done "properly" instead of taking shortcuts commonly taken by humans

# Next Steps

2. Understand the manual artifact (and its components) well



# Next Steps

2. Understand the manual artifact (and its components) well
  - Understanding the problem domain lets one develop a solution that:
    - Makes use of all areas of the domain
    - Follows domain standards, including quality and terminology

# Next Steps

2. Understand the manual artifact (and its components) well
  - Understanding the problem domain lets one develop a solution that:
    - Makes use of all areas of the domain
    - Follows domain standards, including quality and terminology
  - There are specific areas of testing that need to be understood:

# Next Steps

2. Understand the manual artifact (and its components) well
  - Understanding the problem domain lets one develop a solution that:
    - Makes use of all areas of the domain
    - Follows domain standards, including quality and terminology
  - There are specific areas of testing that need to be understood:
    - **Research Question #1:** What information is necessary for different types of testing?

# Next Steps

2. Understand the manual artifact (and its components) well
  - Understanding the problem domain lets one develop a solution that:
    - Makes use of all areas of the domain
    - Follows domain standards, including quality and terminology
  - There are specific areas of testing that need to be understood:
    - **Research Question #1:** What information is necessary for different types of testing?
    - **Research Question #2:** How can test cases be generated from information that currently exists within Drasil?

# Next Steps

2. Understand the manual artifact (and its components) well
  - Understanding the problem domain lets one develop a solution that:
    - Makes use of all areas of the domain
    - Follows domain standards, including quality and terminology
  - There are specific areas of testing that need to be understood:
    - **Research Question #1:** What information is necessary for different types of testing?
    - **Research Question #2:** How can test cases be generated from information that currently exists within Drasil?
    - **Research Question #3:** How can new information be added to facilitate the generation of more types of testing?

# Next Steps

2. Understand the manual artifact (and its components) well
  - Understanding the problem domain lets one develop a solution that:
    - Makes use of all areas of the domain
    - Follows domain standards, including quality and terminology
  - There are specific areas of testing that need to be understood:
    - **Research Question #1:** What information is necessary for different types of testing?
    - **Research Question #2:** How can test cases be generated from information that currently exists within Drasil?
    - **Research Question #3:** How can new information be added to facilitate the generation of more types of testing?

"The information you have should be just as useful for generating tests as it should be for manually running them." —

Dr. Jacques Carette

# Next Steps

3. Generate it!

## 3. Generate it!

- Test cases will then be written for:
  - Other variabilities of Projectile's Python implementation
  - Projectile's implementation in other languages
  - Other examples where code is generated: GlassBR, NoPCM, DbIPendulum, PD Controller (Hunt et al., 2021)



## 3. Generate it!

- Test cases will then be written for:
  - Other variabilities of Projectile's Python implementation
  - Projectile's implementation in other languages
  - Other examples where code is generated: GlassBR, NoPCM, DbIPendulum, PD Controller (Hunt et al., 2021)
- These test cases will also be added to Drasil's CI/CD to ensure that future changes preserve the code's functionality

# Acknowledgment

- Dr. Smith and Dr. Carette have been great supervisors in the past and have, both then and now, provided me with valuable guidance and feedback
  - They have helped me refine the scope of this project
  - The project itself was originally posed by Dr. Smith back in 2020!

# Acknowledgment

- Dr. Smith and Dr. Carette have been great supervisors in the past and have, both then and now, provided me with valuable guidance and feedback
  - They have helped me refine the scope of this project
  - The project itself was originally posed by Dr. Smith back in 2020!
- The format of this presentation was *heavily* based on a previous presentation by Jason Balaci, who also provided a great thesis template

# Acknowledgment

- Dr. Smith and Dr. Carette have been great supervisors in the past and have, both then and now, provided me with valuable guidance and feedback
  - They have helped me refine the scope of this project
  - The project itself was originally posed by Dr. Smith back in 2020!
- The format of this presentation was *heavily* based on a previous presentation by Jason Balaci, who also provided a great thesis template
- The past and current Drasil team have created a truly amazing framework!

Thank you!  
Questions?

# Table of Contents

## 1 Introduction

- The Need for Standardized Terminology
- The Lack of Standardized Terminology

## 2 Project

- Drasil
- Why Test Generated Code?
- Next Steps

## 3 References

# References

- Matthias Hamburg and Gary Mogyorodi, editors. ISTQB Glossary, v4.3, 2024. URL [https://glossary.istqb.org/en\\_US/search](https://glossary.istqb.org/en_US/search).
- Anthony Hunt, Peter Michalski, Dong Chen, Jason Balaci, and Spencer Smith. Drasil - Generate All the Things!, 2021. URL <https://jacquescurette.github.io/Drasil/>.
- ISO/IEC and IEEE. ISO/IEC/IEEE International Standard - Systems and software engineering—Vocabulary. *ISO/IEC/IEEE 24765:2017(E)*, September 2017. doi: 10.1109/IEEESTD.2017.8016712.
- ISO/IEC and IEEE. ISO/IEC/IEEE International Standard - Systems and software engineering—Software testing—Part 1: General concepts. *ISO/IEC/IEEE 29119-1:2022(E)*, January 2022. doi: 10.1109/IEEESTD.2022.9698145.
- Cem Kaner, James Bach, and Bret Pettichord. *Lessons Learned in Software Testing: A Context-Driven Approach*. John Wiley & Sons, December 2011. ISBN 978-0-471-08112-8. URL <https://www.wiley.com/en-ca/Lessons+Learned+in+Software+Testing%3A+A+Context-Driven+Approach-p-9780471081128>. 