# Empirically Evaluating Regression Testing Techniques: Challenges, Solutions, and a Potential Way Forward

Gregory M. Kapfhammer
Department of Computer Science
Allegheny College
http://www.cs.allegheny.edu/~gkapfham/


ALLEGHENY COLLEGE

1st International Workshop on Regression Testing
Co-Located with the 4th IEEE International Conference on Software Testing, Verification and Validation
Berlin, Germany, March 2011

# Presentation Overview

Regression Testing Community

# Presentation Overview

Current Trends

Regression Testing Community

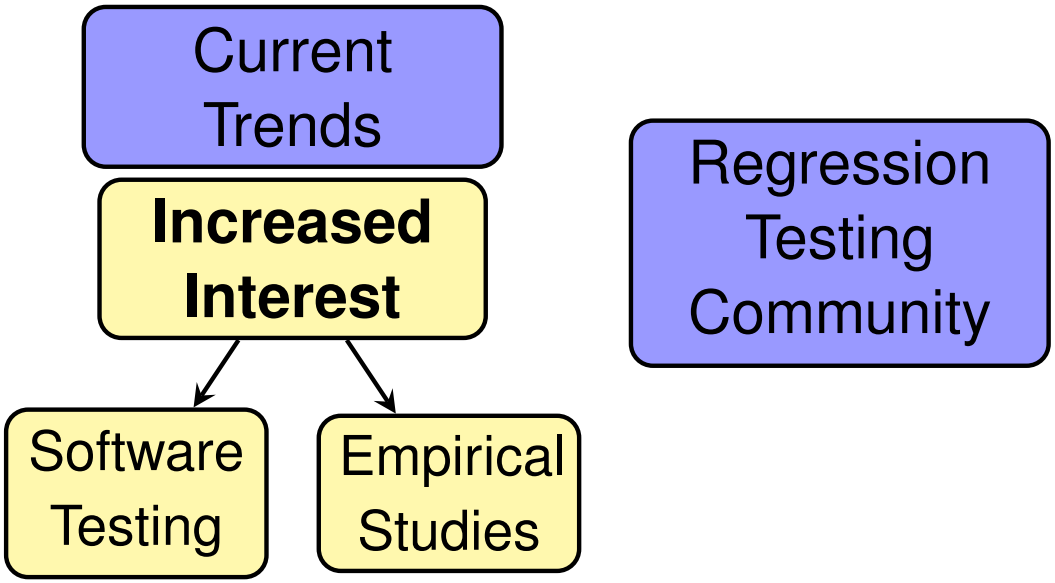
# Presentation Overview

Current Trends

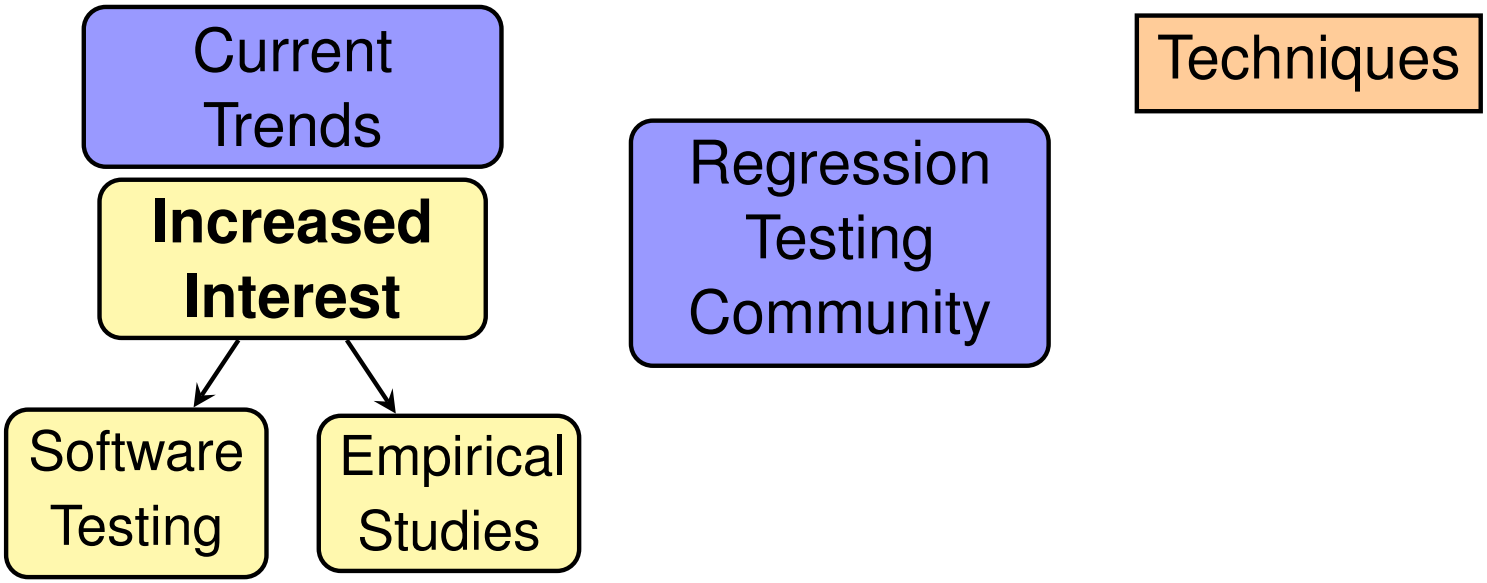
**Increased Interest**

Regression Testing Community

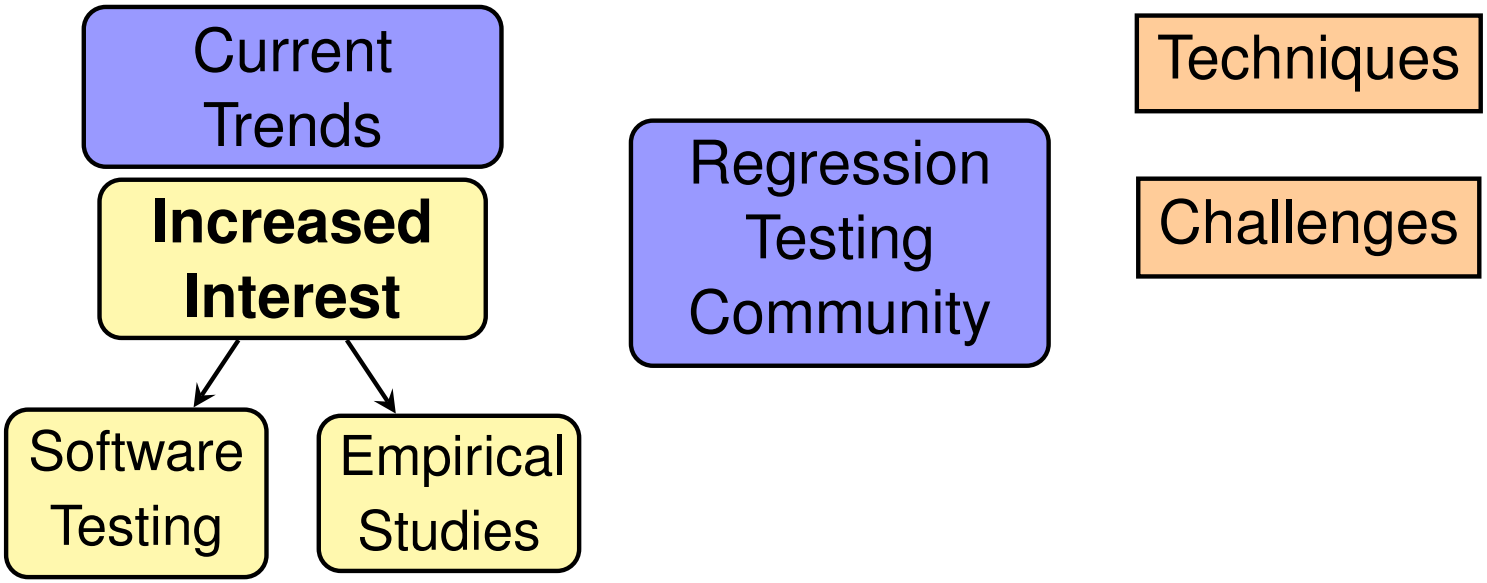
# Presentation Overview



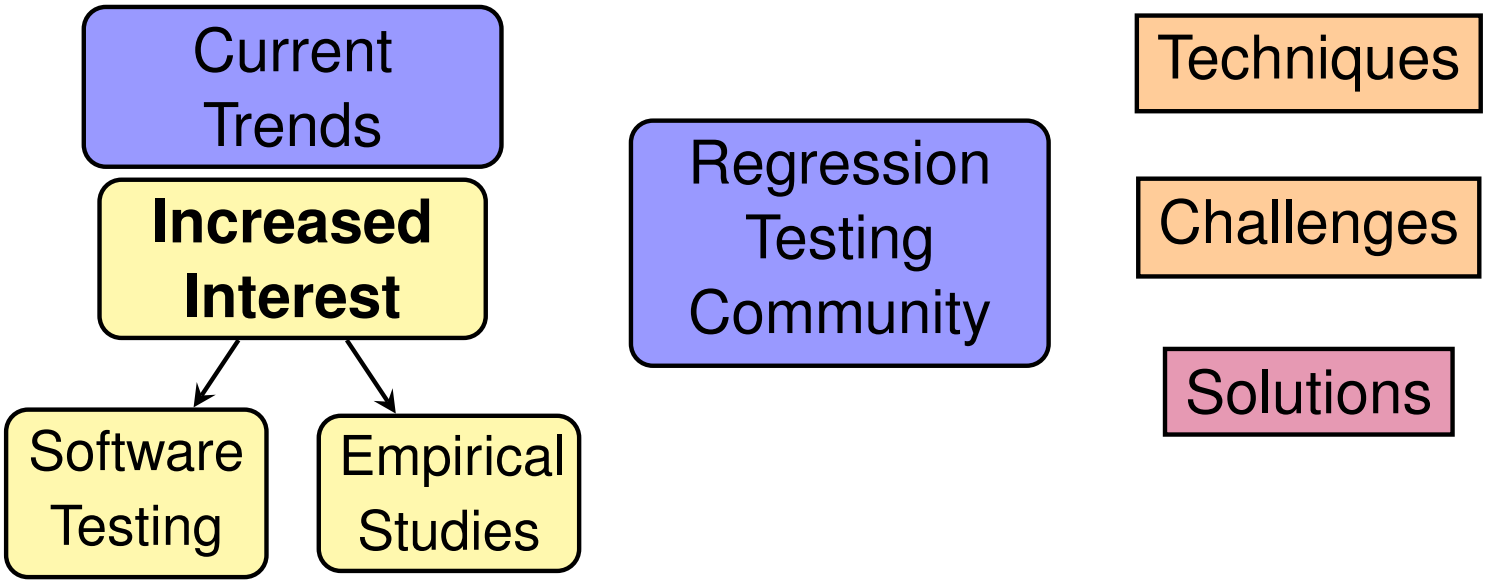
# Presentation Overview



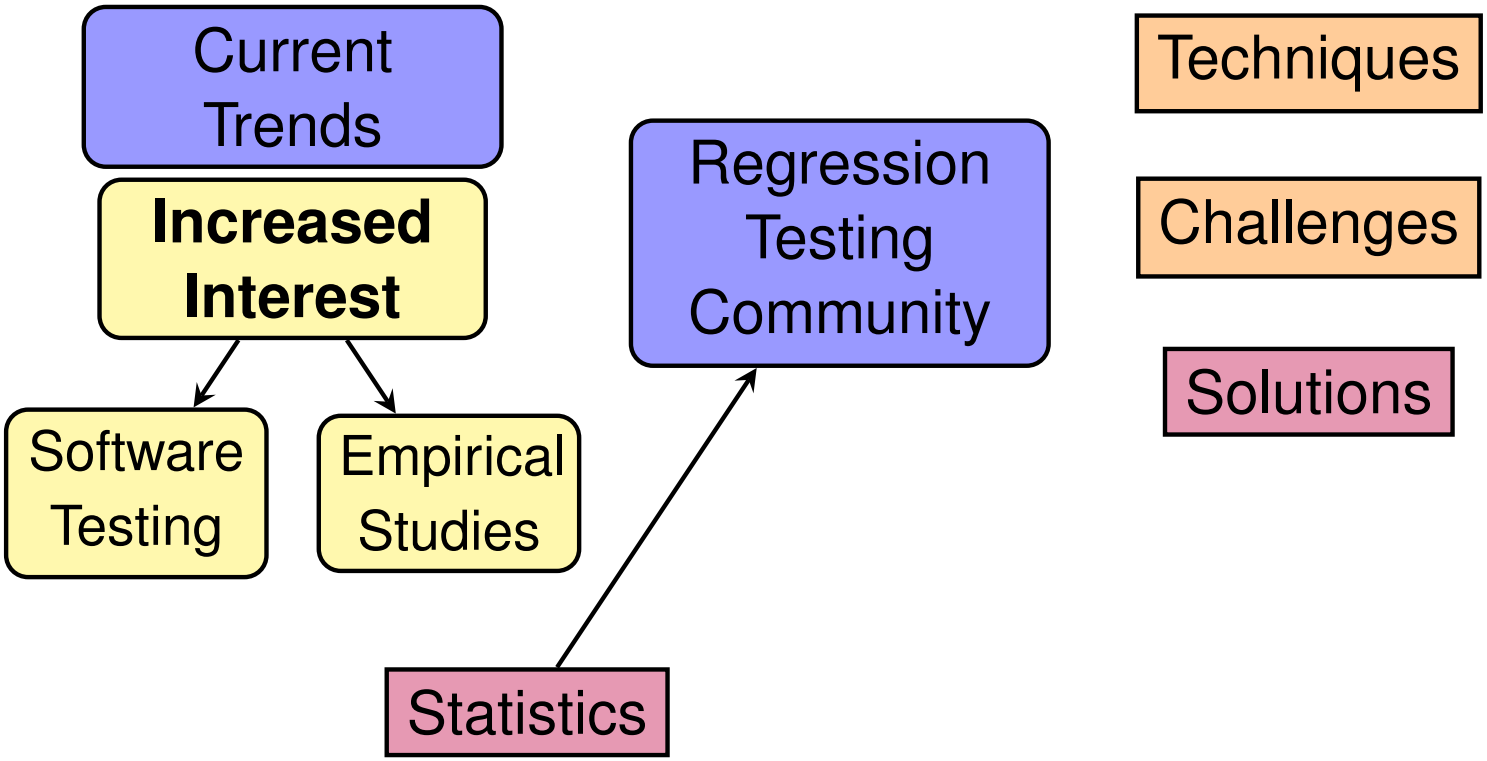
# Presentation Overview



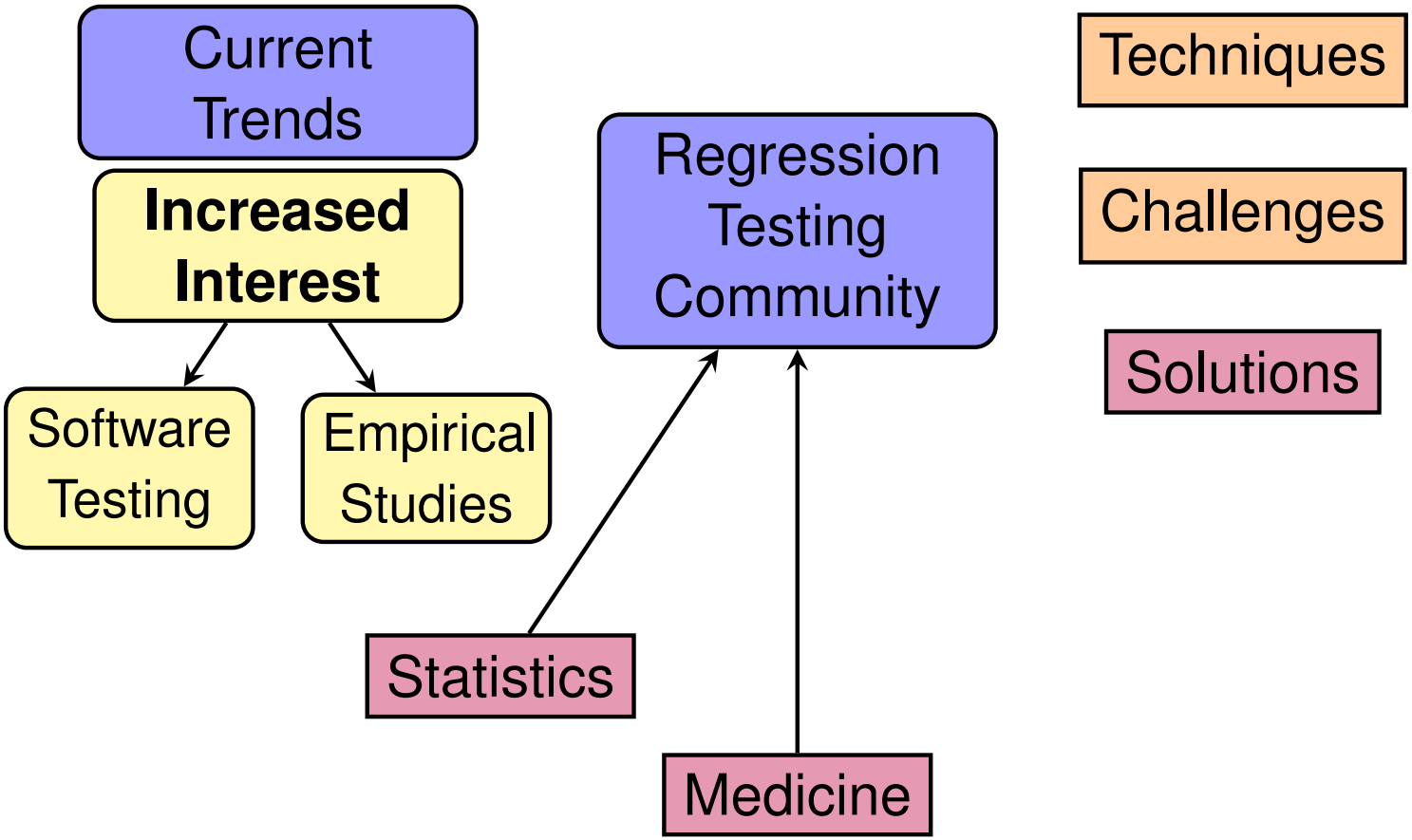
# Presentation Overview



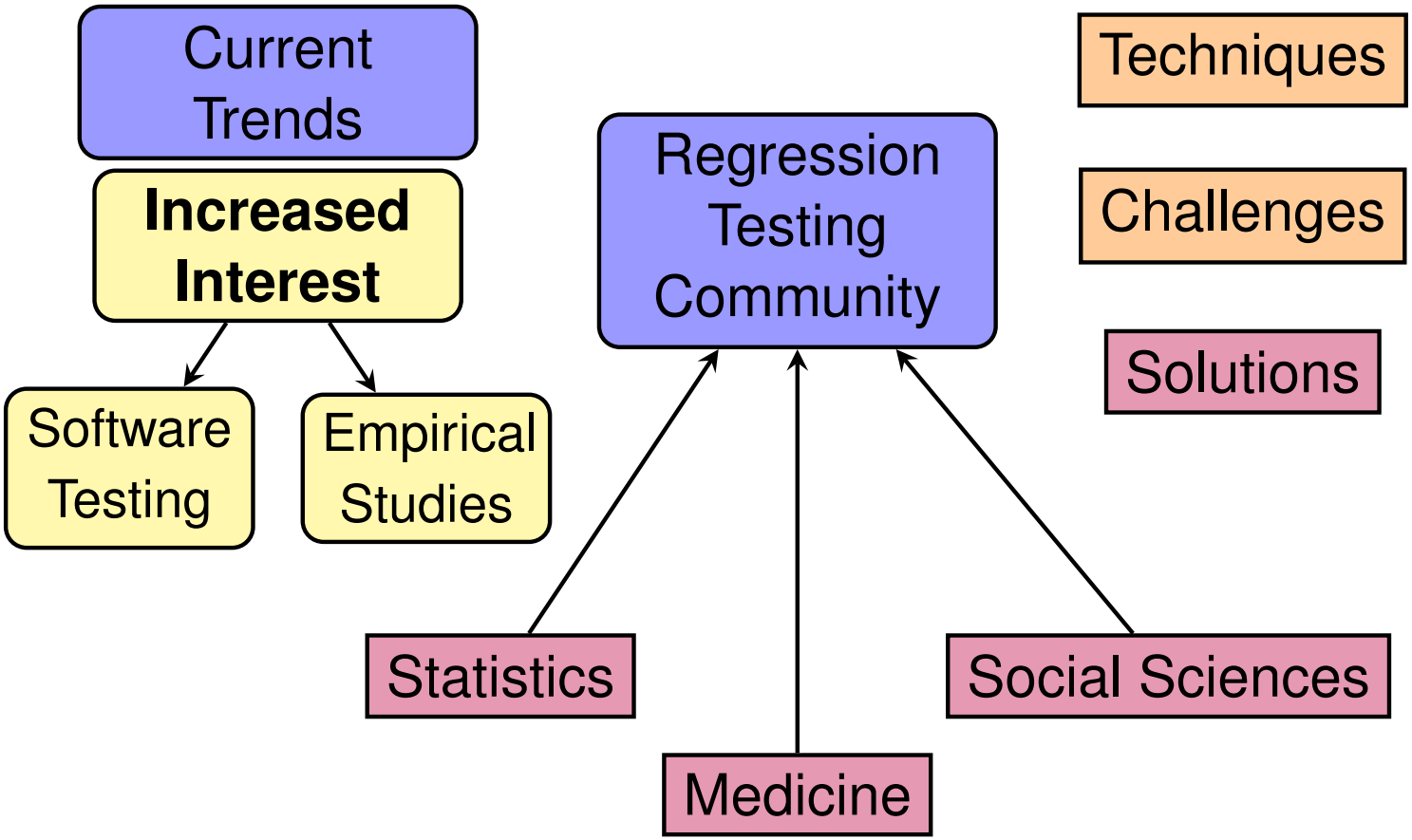
# Presentation Overview



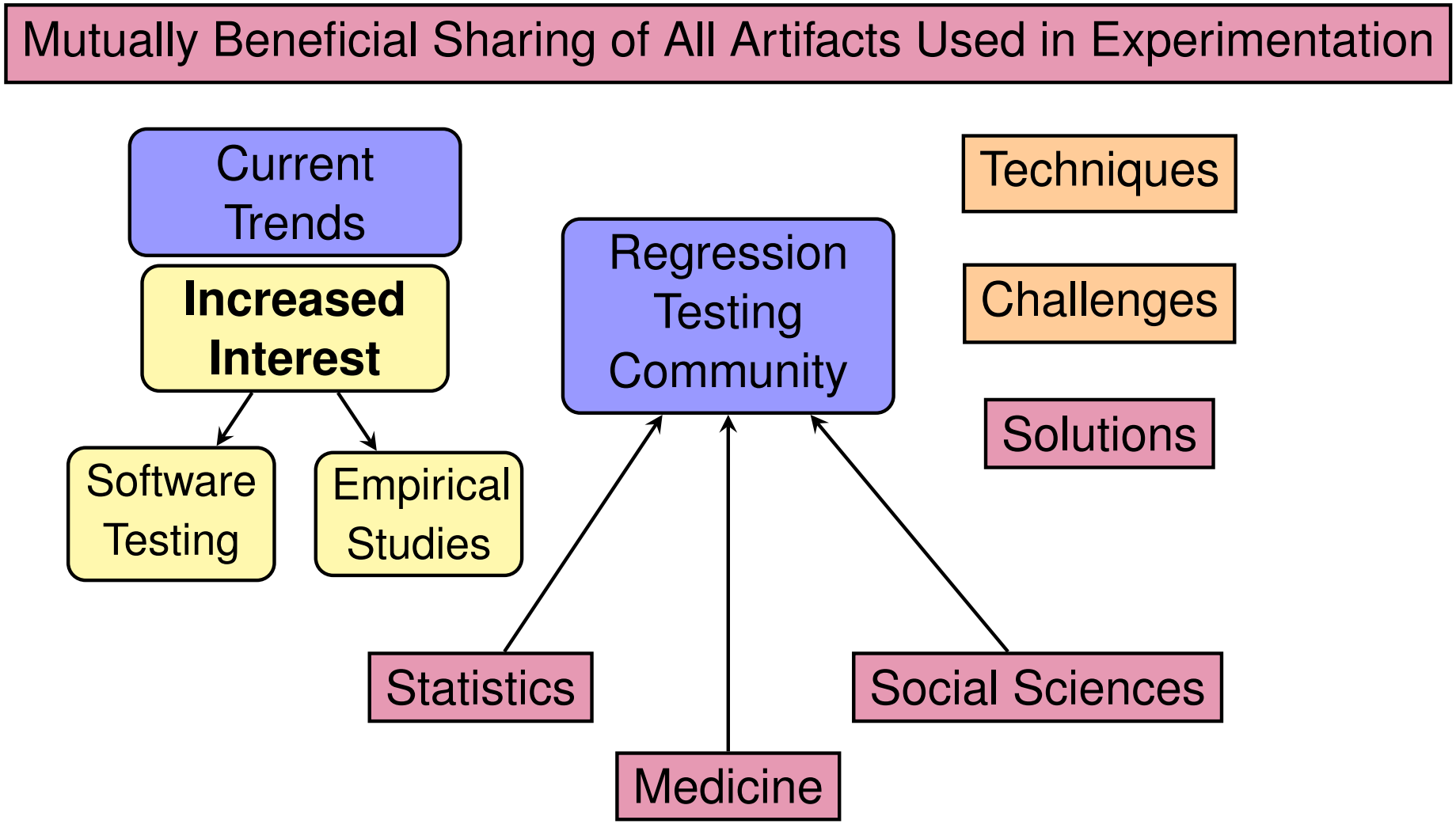
# Presentation Overview



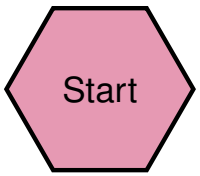
# Presentation Overview



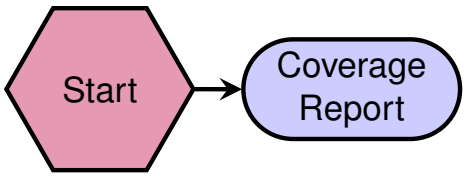
# Presentation Overview



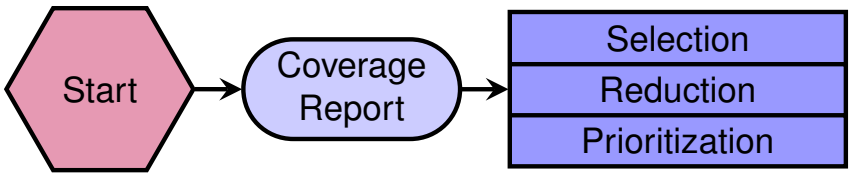
# Model of Regression Testing



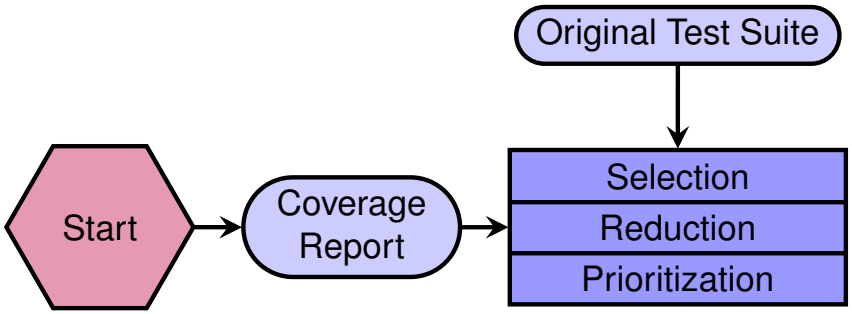
# Model of Regression Testing



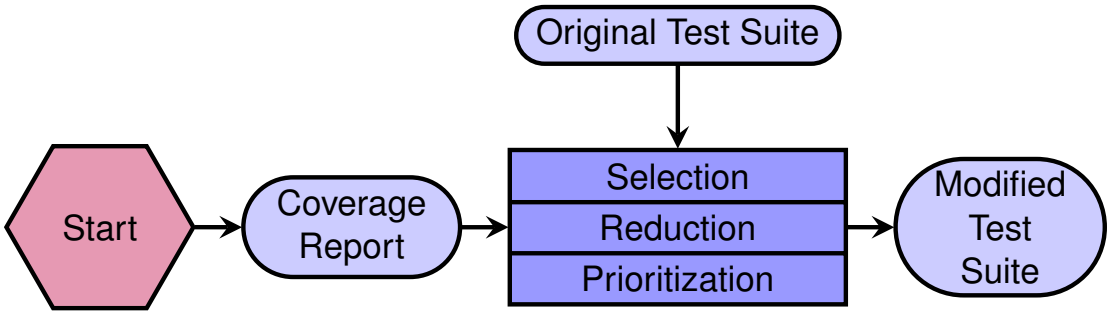
# Model of Regression Testing



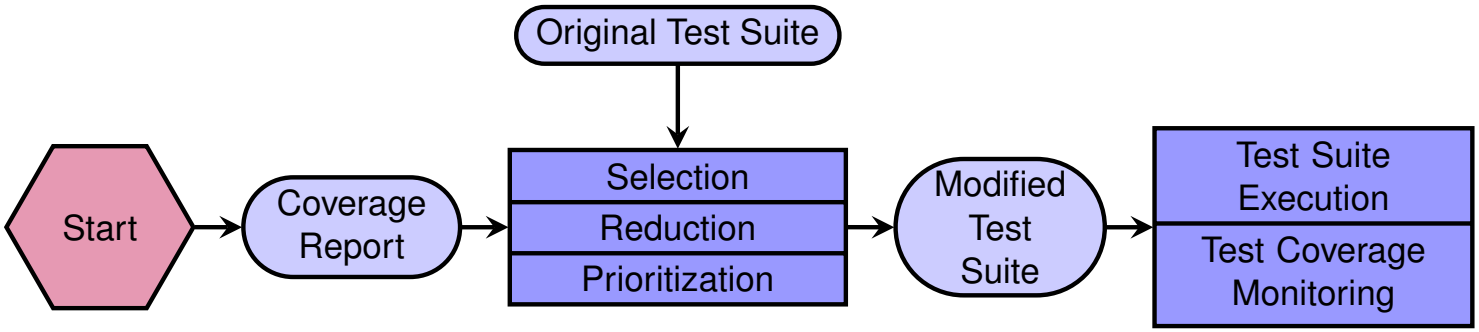
# Model of Regression Testing



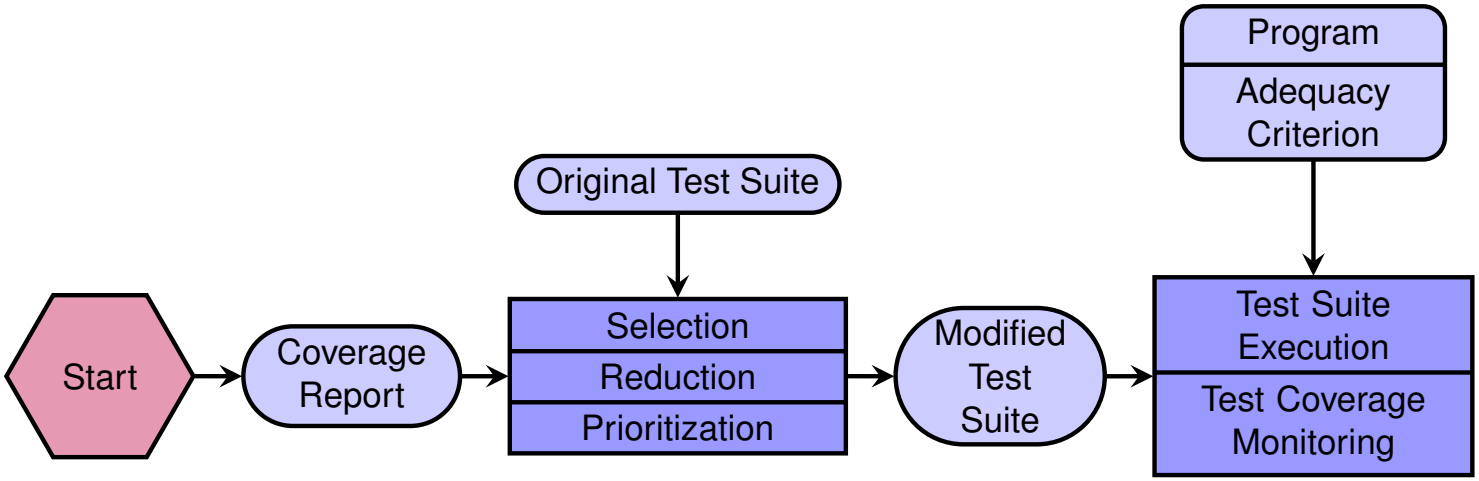
# Model of Regression Testing



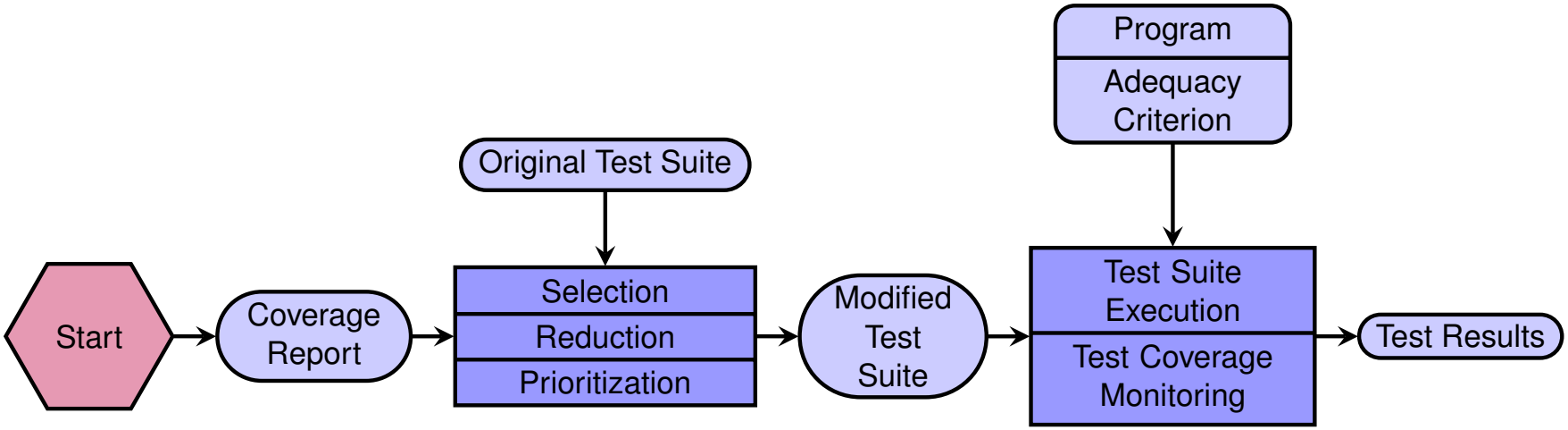
# Model of Regression Testing



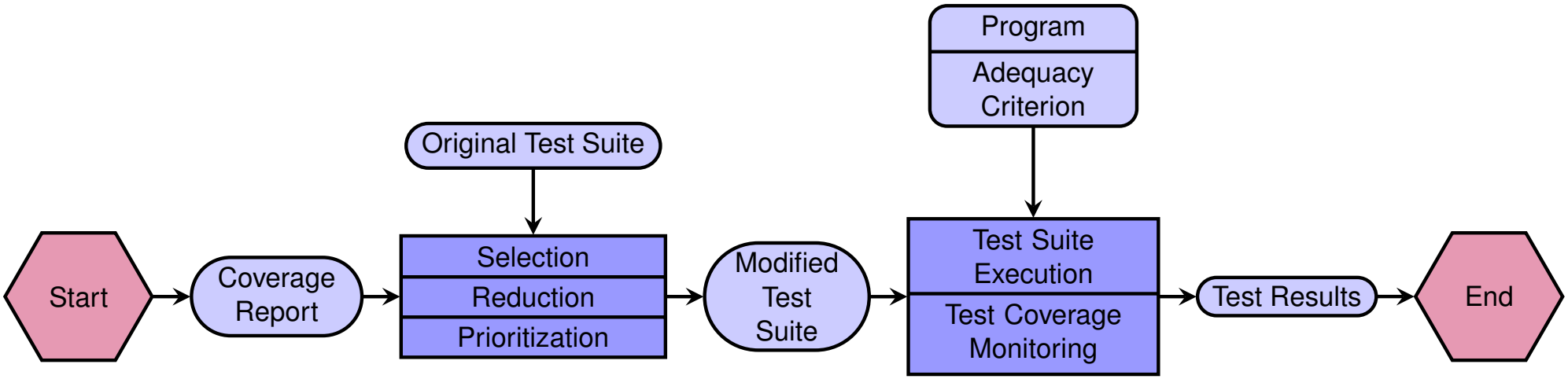
# Model of Regression Testing



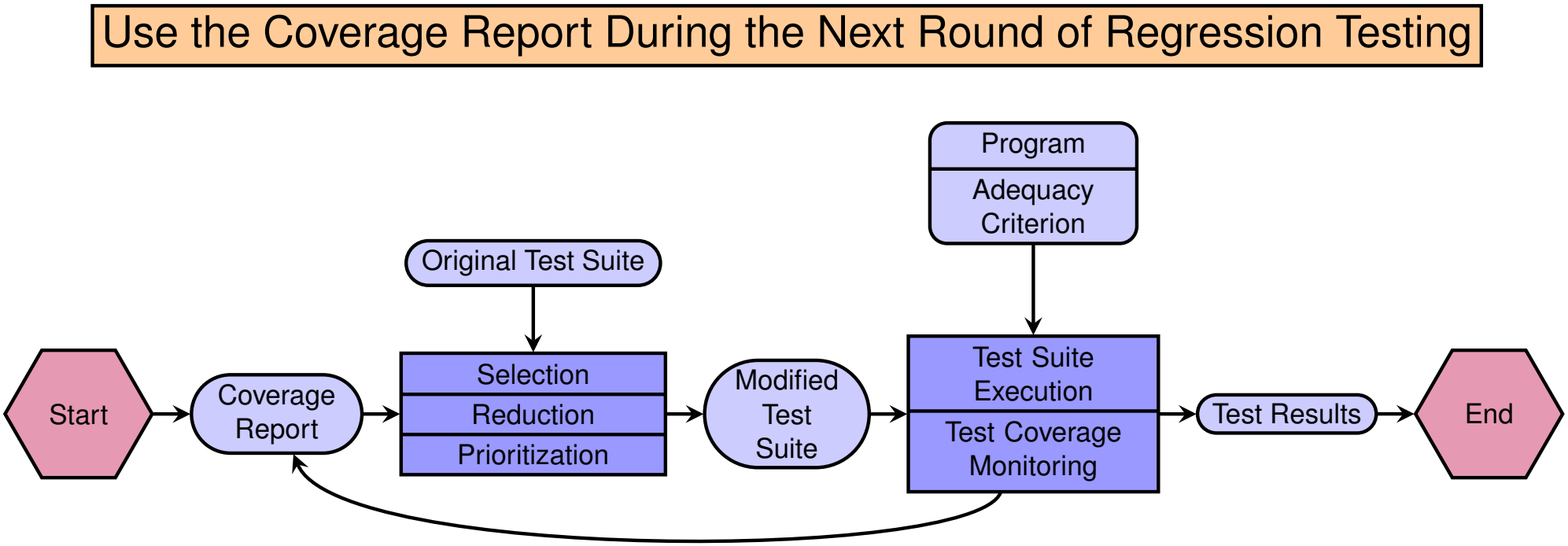
# Model of Regression Testing



# Model of Regression Testing

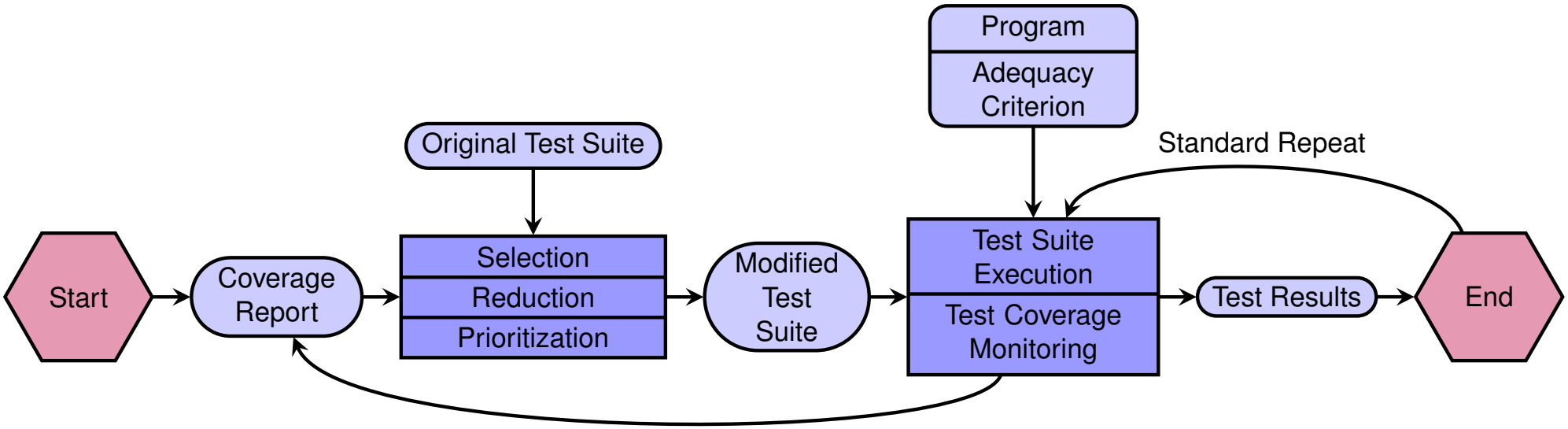


# Model of Regression Testing

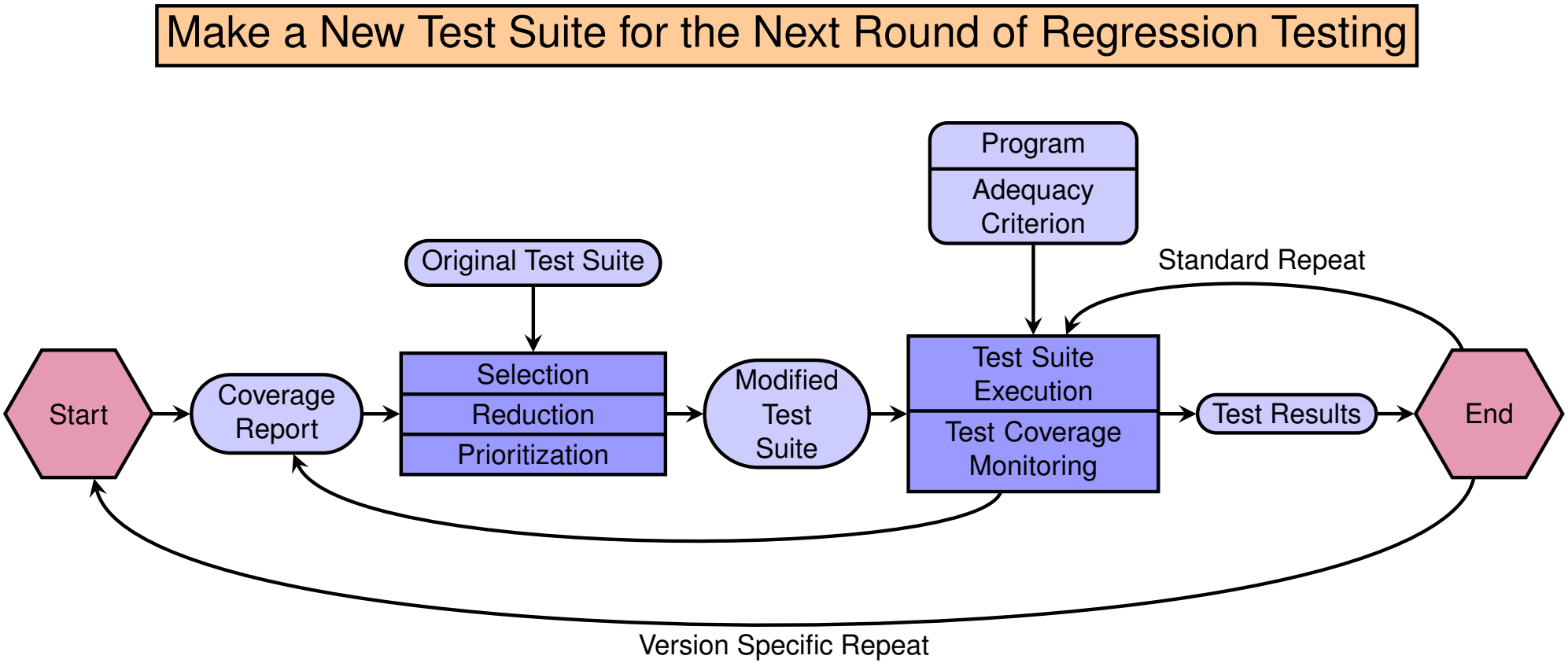


# Model of Regression Testing

Use the Same Test Suite for the Next Round of Regression Testing

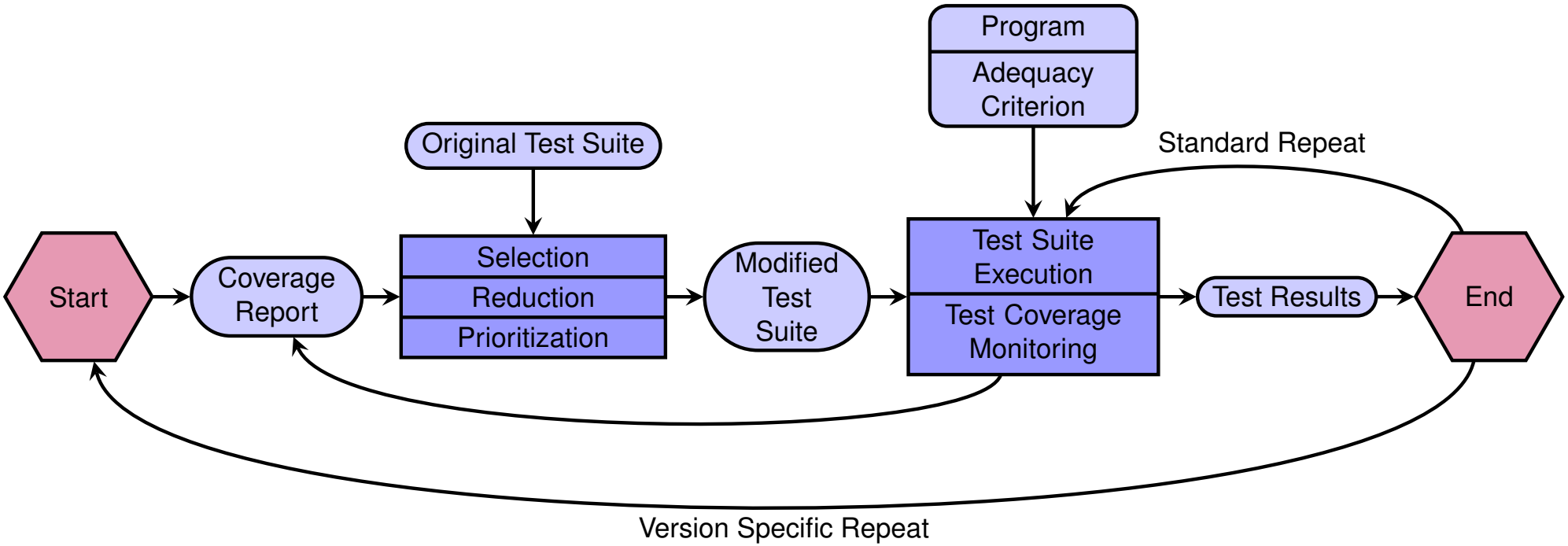


# Model of Regression Testing



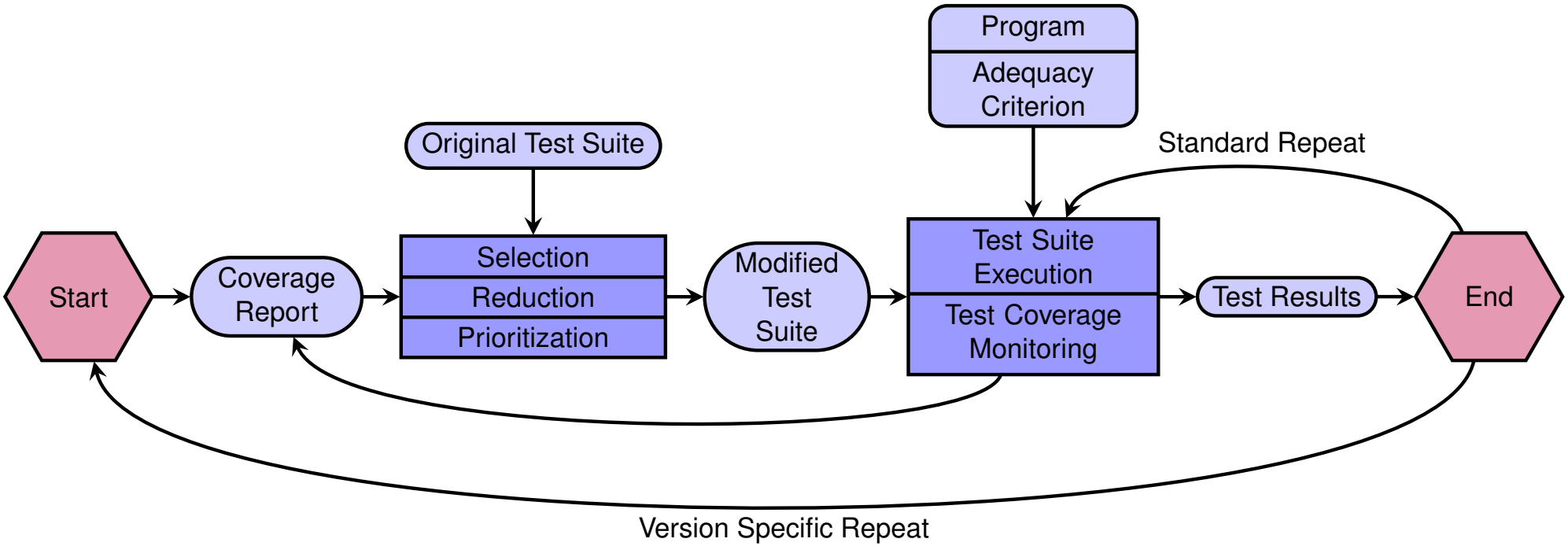
# Model of Regression Testing

Practitioners are unwilling to use methods for “lack of empirical studies” [12]



# Model of Regression Testing

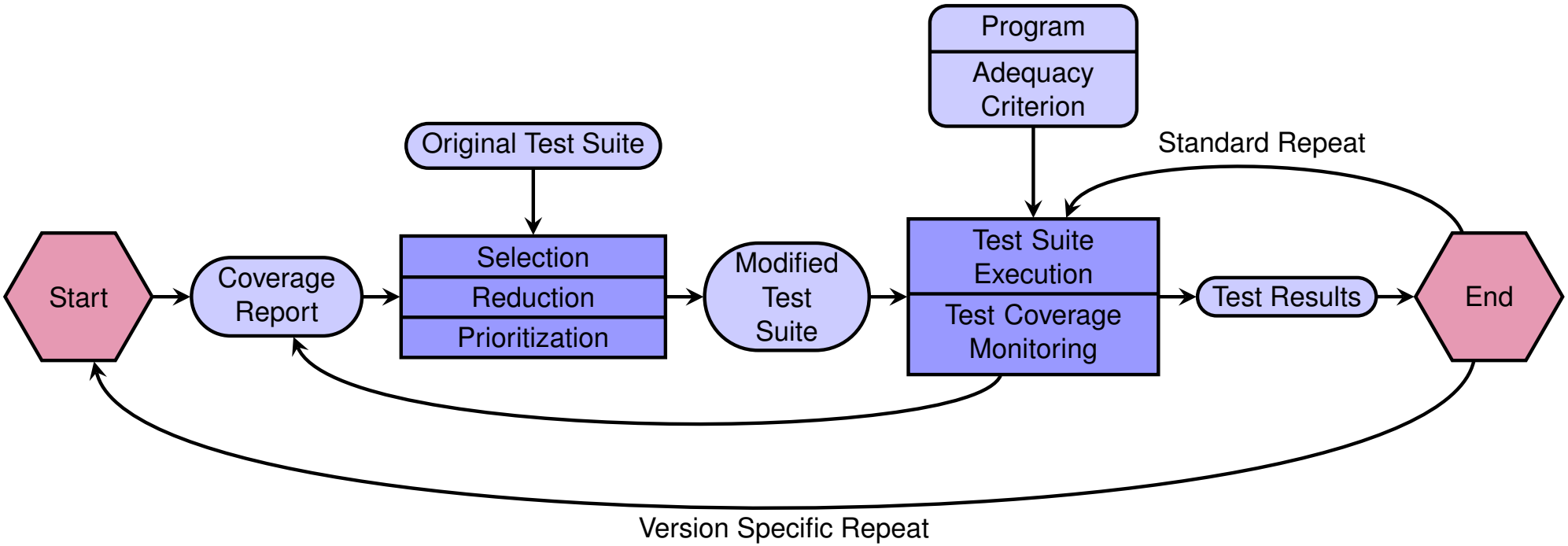
Practitioners are unwilling to use methods for “lack of empirical studies” [12]




Testing tools do not produce outputs in the best format (e.g., per-test coverage)

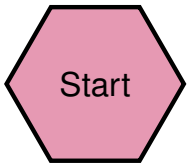
# Model of Regression Testing

Practitioners are unwilling to use methods for “lack of empirical studies” [12]

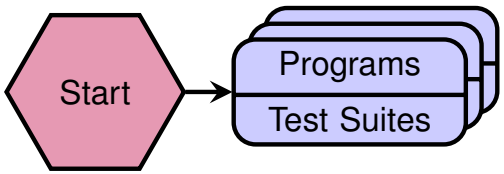



Experimental assessment could stagnate due to inaccessibility of artifacts

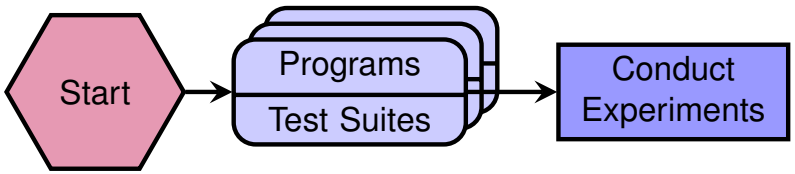
# Model of Experimental Evaluation



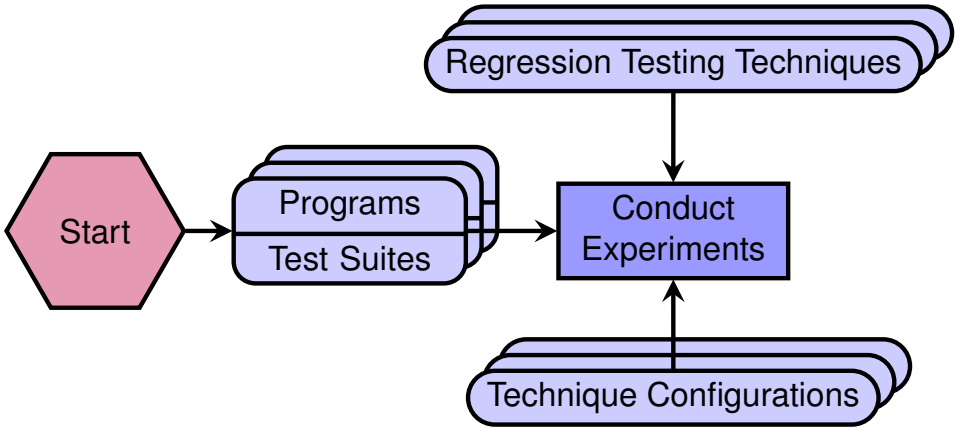
# Model of Experimental Evaluation



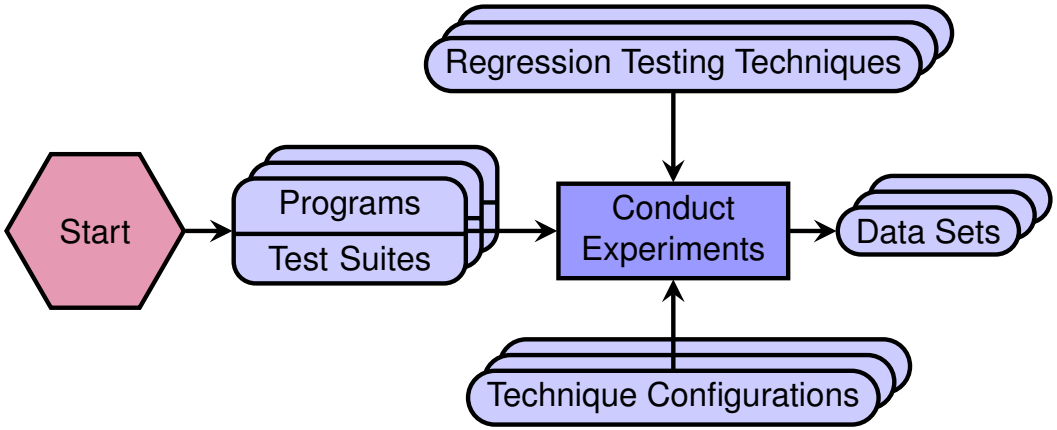
# Model of Experimental Evaluation



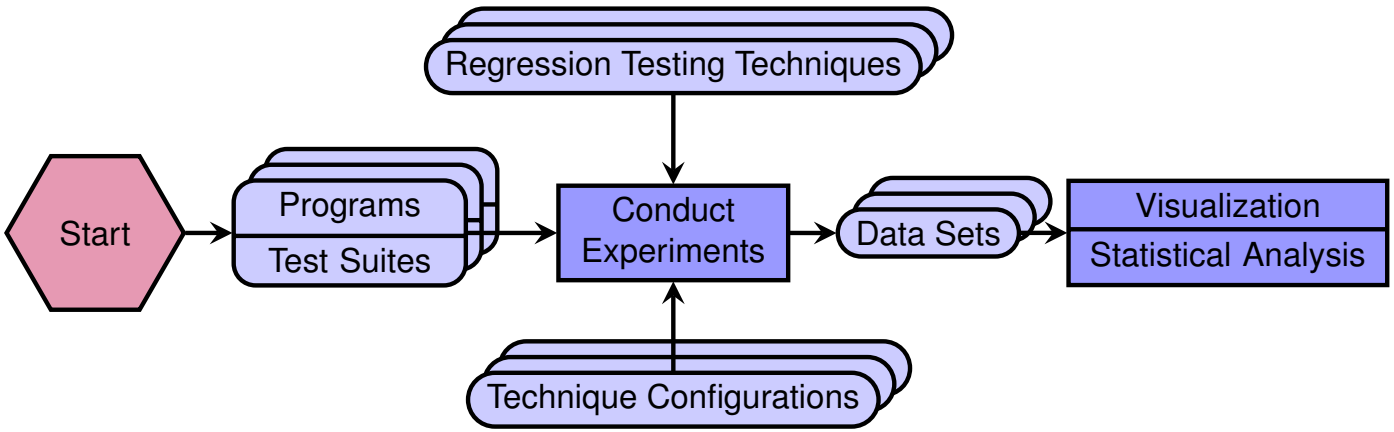
# Model of Experimental Evaluation



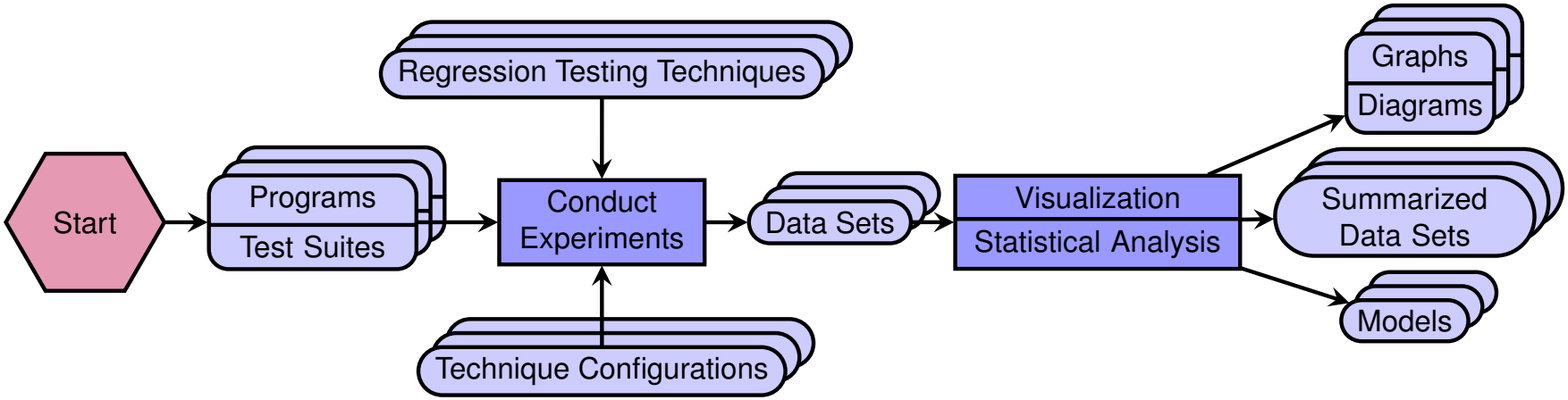
# Model of Experimental Evaluation



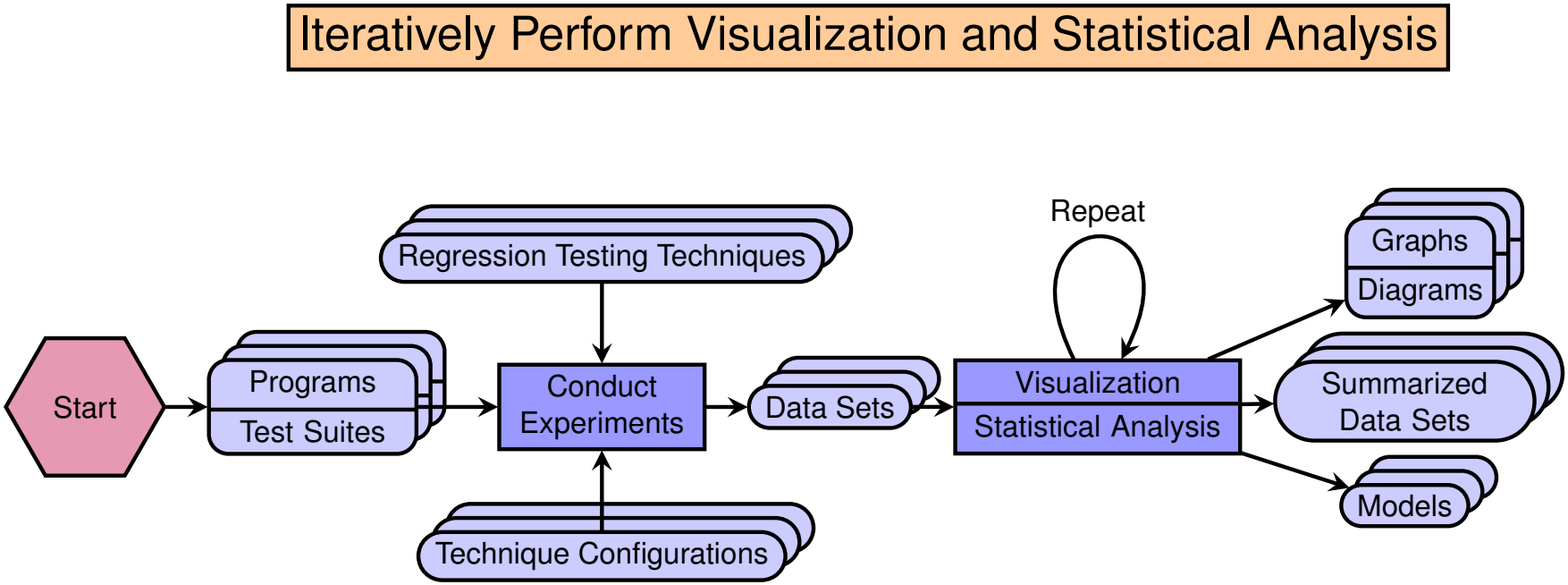
# Model of Experimental Evaluation



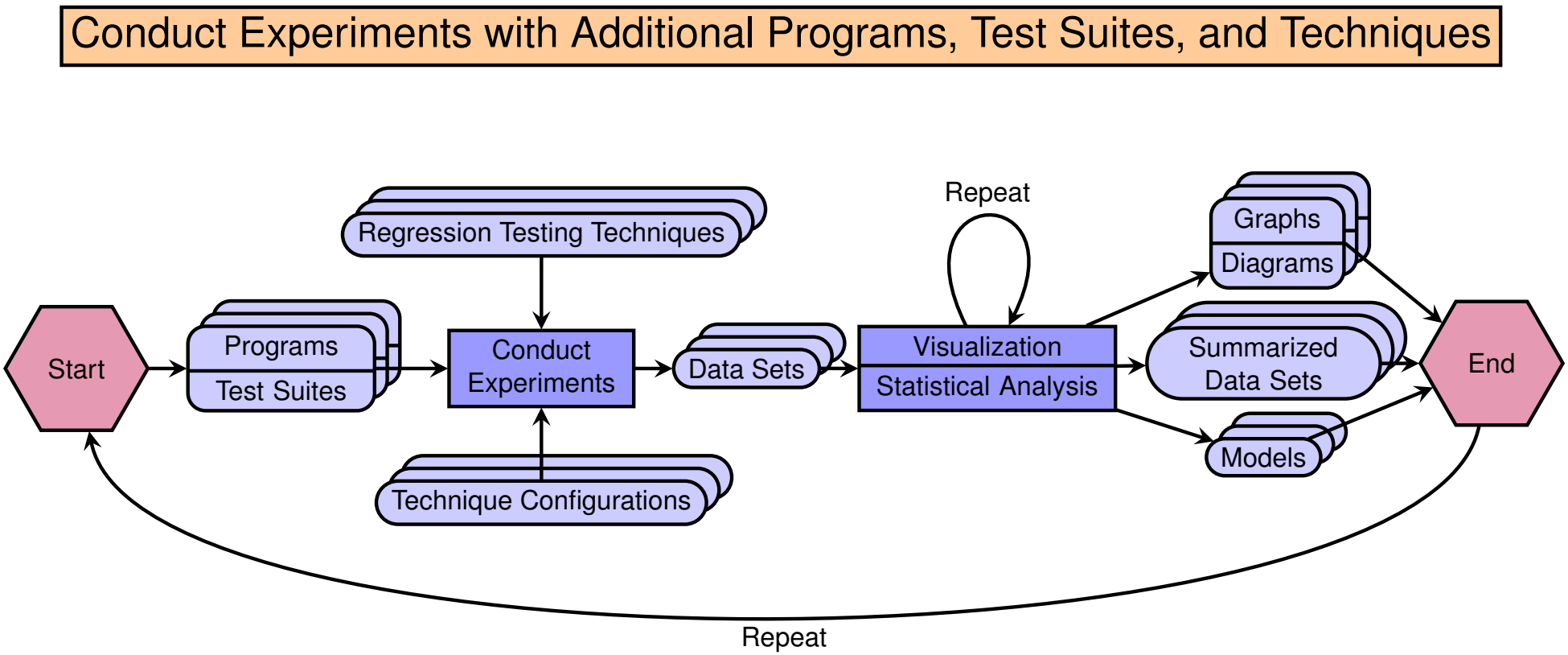
# Model of Experimental Evaluation



# Model of Experimental Evaluation

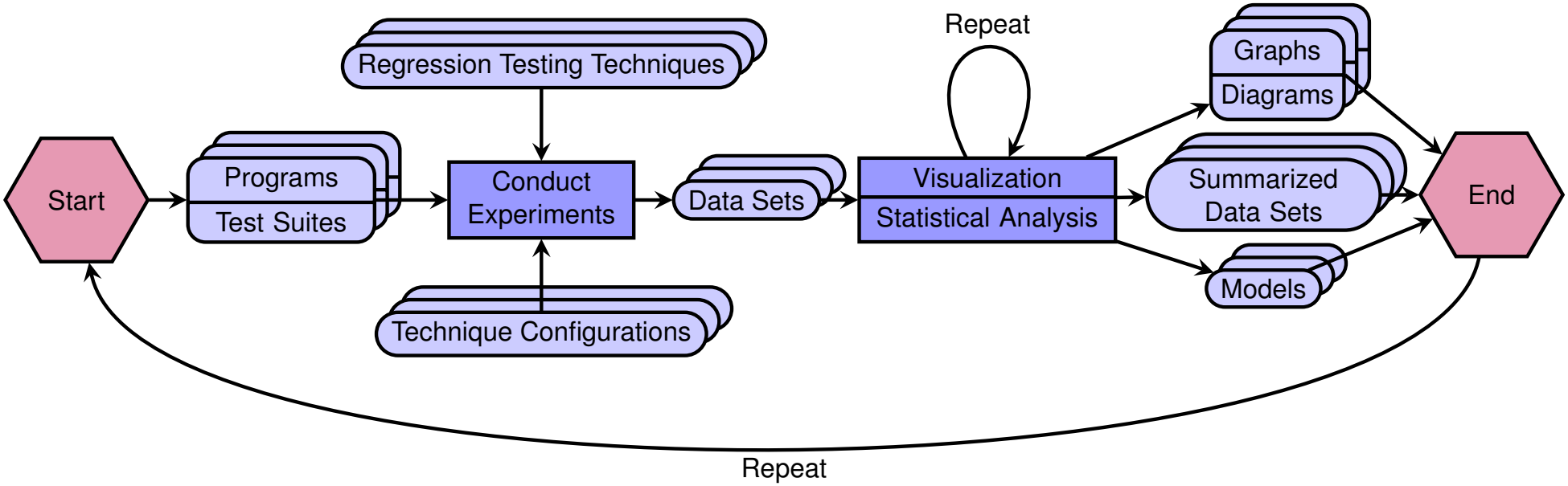


# Model of Experimental Evaluation



# Model of Experimental Evaluation

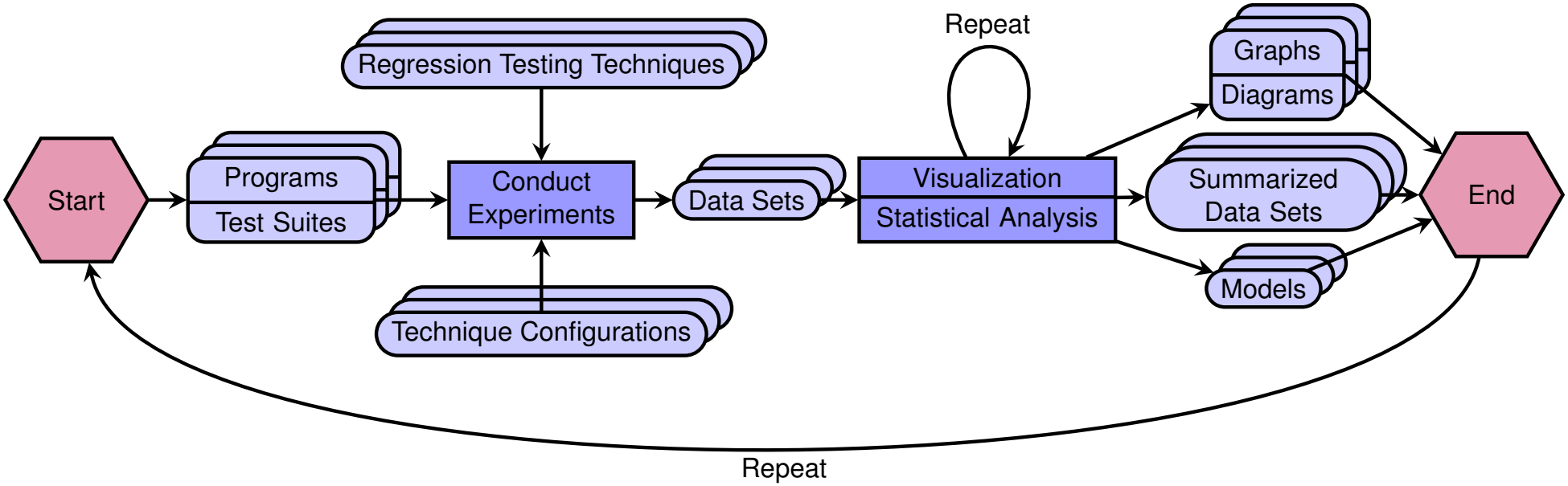
Conduct Experiments with Additional Programs, Test Suites, and Techniques




56% of Papers Surveyed by Yoo and Harman Only Used the SIR Programs [3]

# Model of Experimental Evaluation

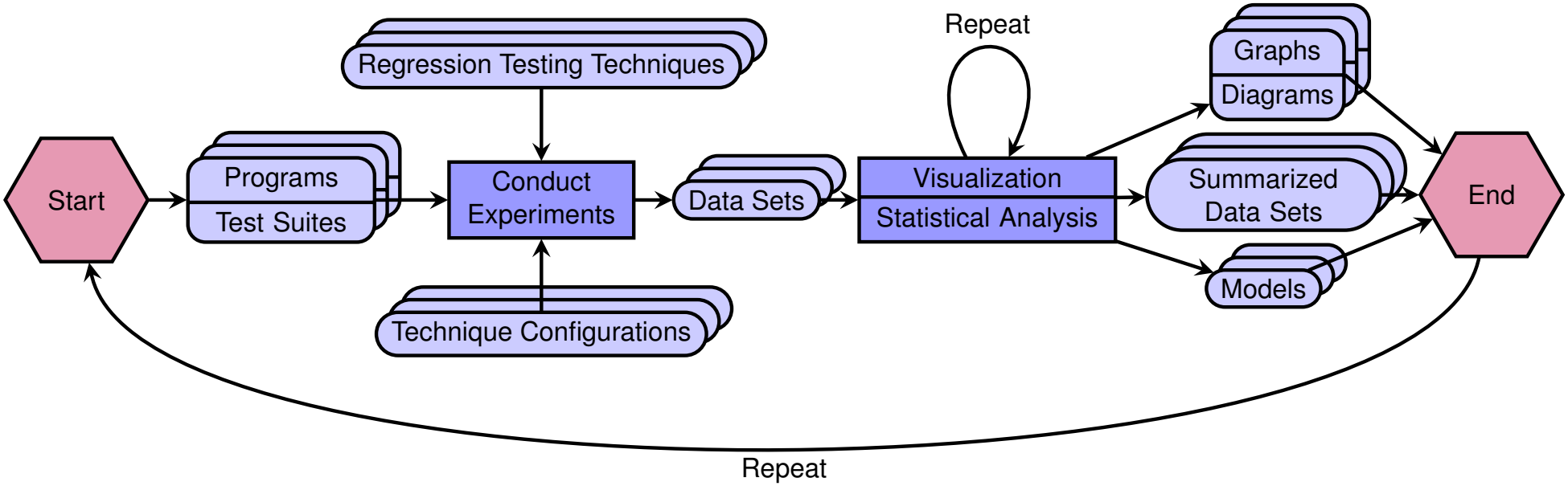
Conduct Experiments with Additional Programs, Test Suites, and Techniques




Few Papers Report on the Efficiency of the Regression Testing Techniques

# Model of Experimental Evaluation

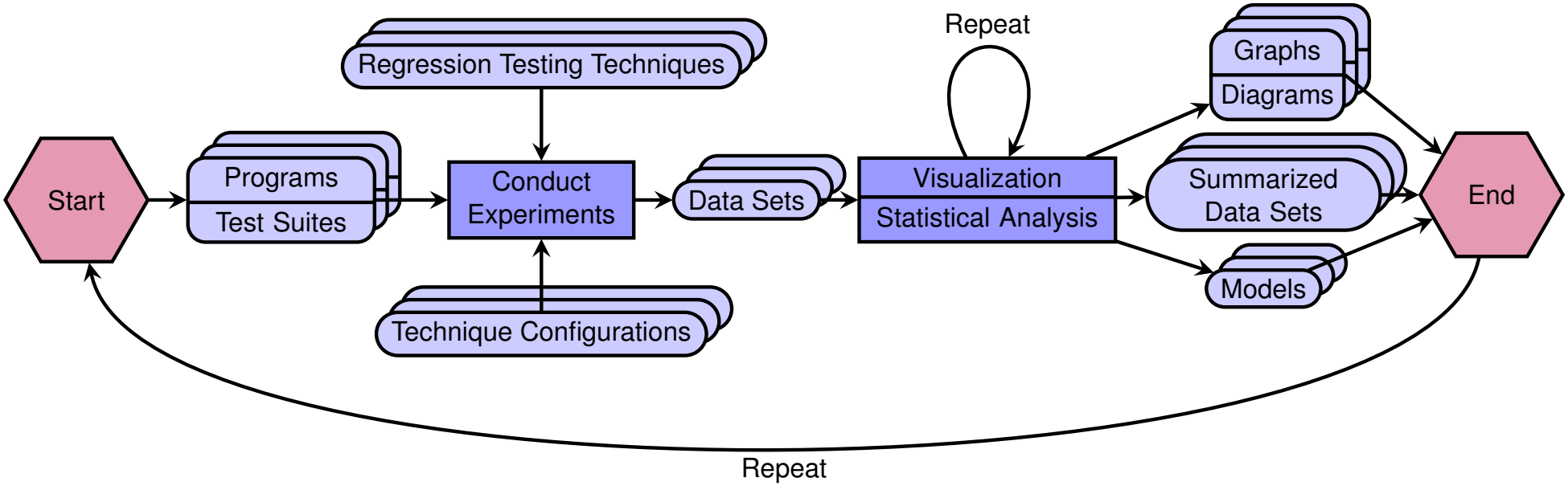
Conduct Experiments with Additional Programs, Test Suites, and Techniques




Authors Do Not Release Tools That Conduct Experiments and Analyze Results

# Model of Experimental Evaluation

Conduct Experiments with Additional Programs, Test Suites, and Techniques




Without Using Data Mining Methods, Researchers May Miss Important Trends

# Thought Provoking Assertions

Jonathan B. Buckheit
and David L. Donoho
Department of Statistics
Stanford University

# Thought Provoking Assertions

For a field to qualify as a science, it is important first and foremost that published work be reproducible by others.

Jonathan B. Buckheit
and David L. Donoho
Department of Statistics
Stanford University

# Thought Provoking Assertions

For a field to qualify as a science, it is important first and foremost that published work be reproducible by others.

Does the Regression Testing Community Want to be Scientific?

Jonathan B. Buckheit and David L. Donoho
Department of Statistics
Stanford University

# Thought Provoking Assertions

For a field to qualify as a science, it is important first and foremost that published work be reproducible by others.

Does the Regression Testing Community Want to be Scientific?

What Does it Mean for Research to be Reproducible?

Jonathan B. Buckheit and David L. Donoho
Department of Statistics
Stanford University

# Thought Provoking Assertions

Gary King
Department of Government Harvard University

# Thought Provoking Assertions

**Reproducible**: sufficient information exists with which to understand, evaluate, and build upon a prior work if a third party can replicate the results without any additional information from the author.

Gary King
Department of Government Harvard University

# Thought Provoking Assertions

**Reproducible**: sufficient information exists with which to understand, evaluate, and build upon a prior work if a third party can replicate the results without any additional information from the author.

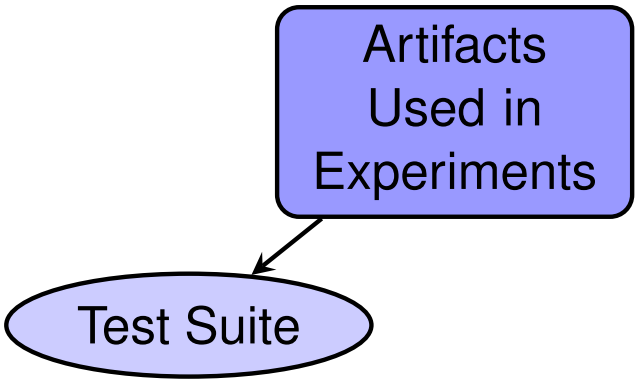
Can the Regression Testing Community Adhere to This Standard?

Gary King
Department of Government Harvard University

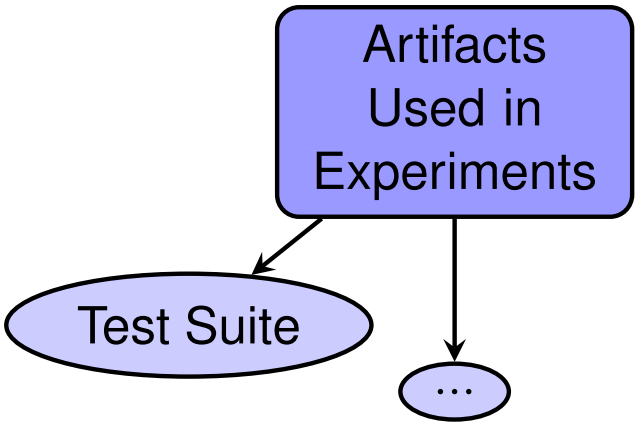
# Reproducible Research Through Sharing

Artifacts Used in Experiments

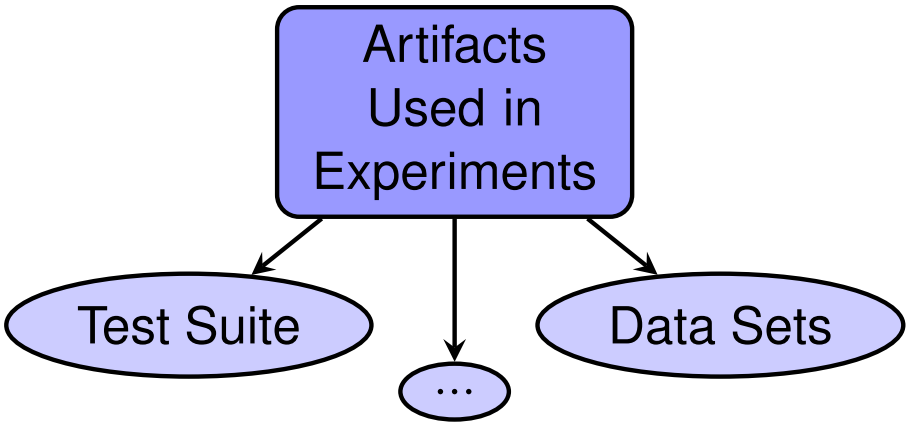
# Reproducible Research Through Sharing



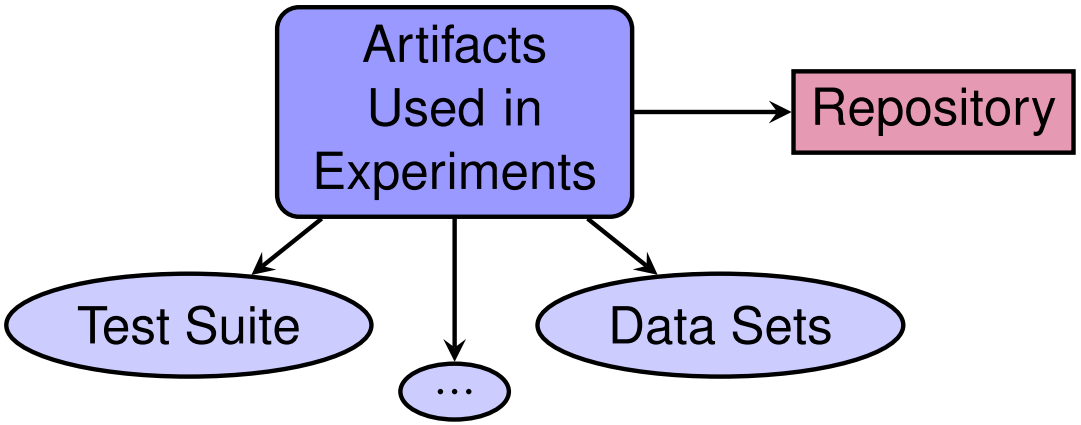
# Reproducible Research Through Sharing



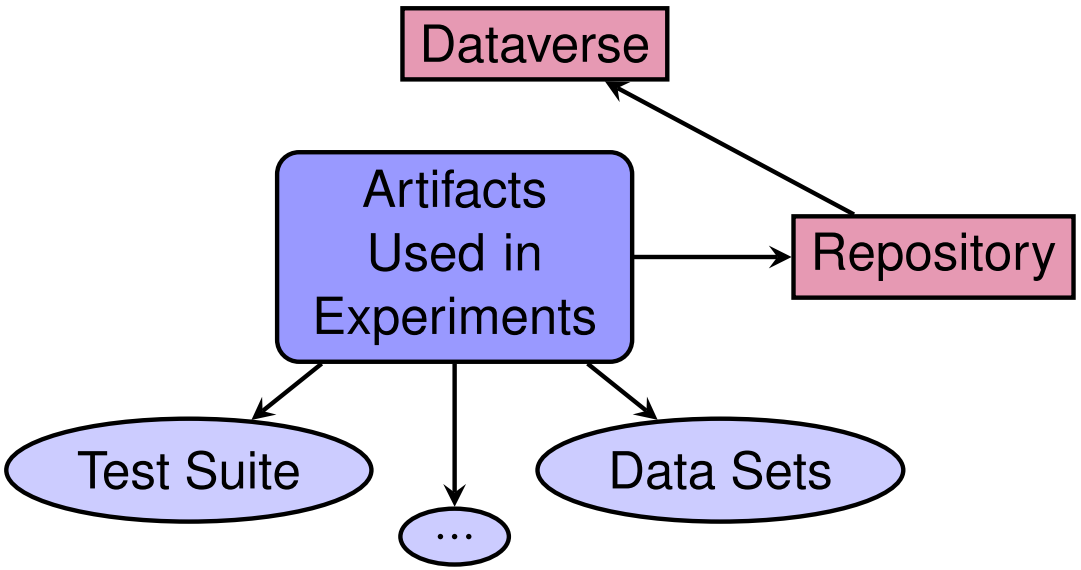
# Reproducible Research Through Sharing



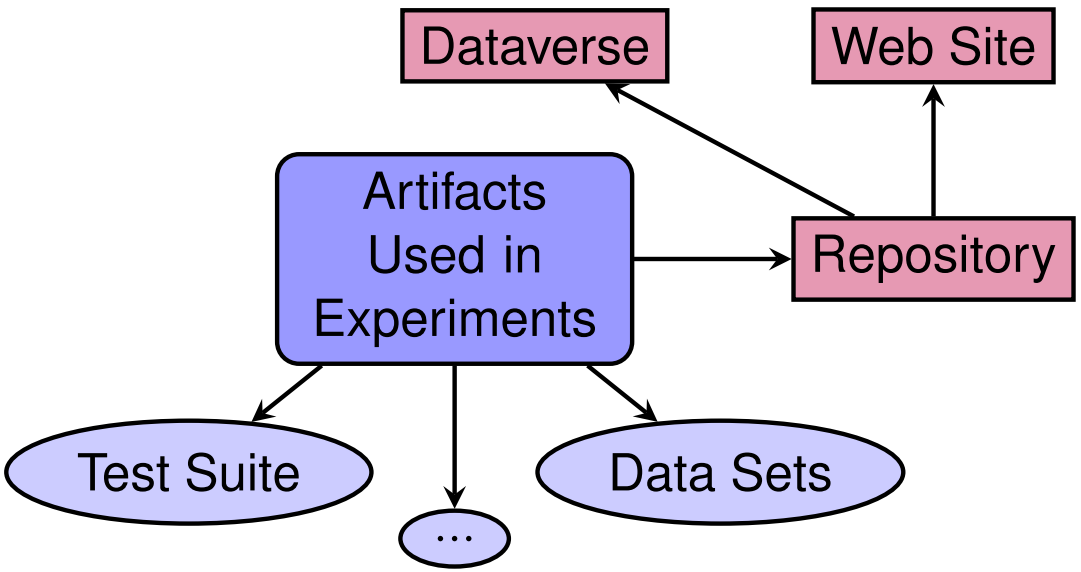
# Reproducible Research Through Sharing



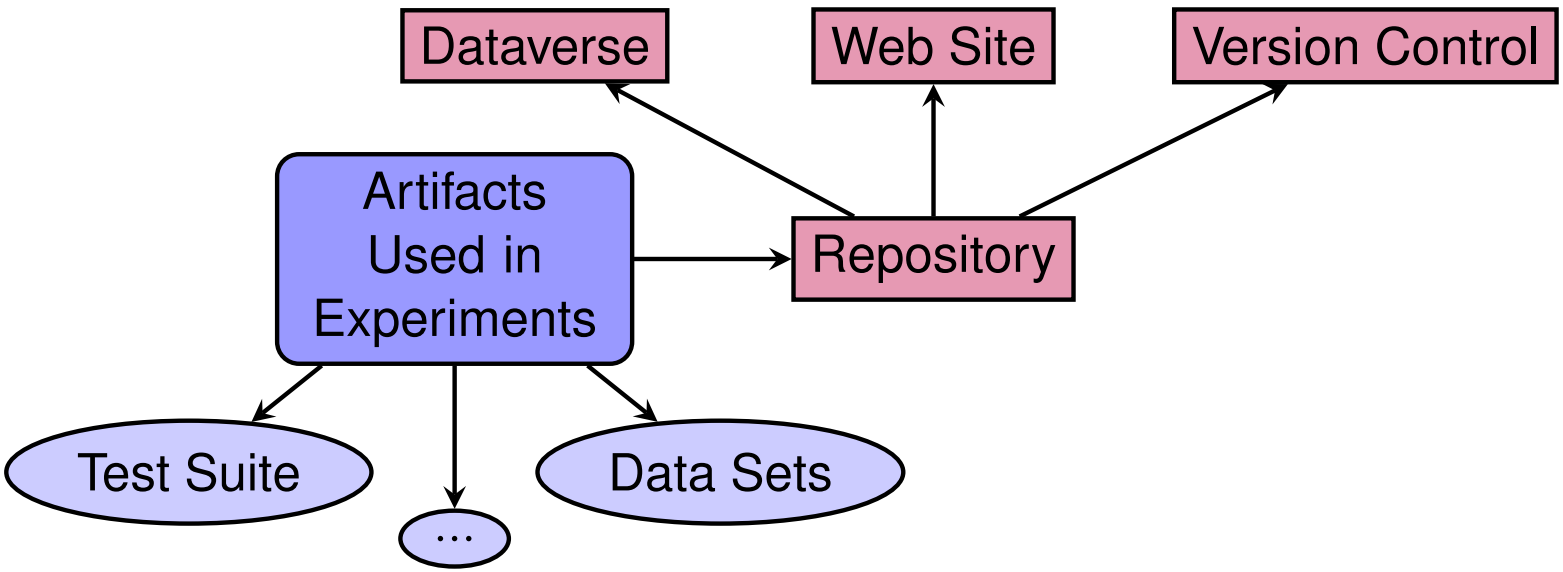
# Reproducible Research Through Sharing



# Reproducible Research Through Sharing

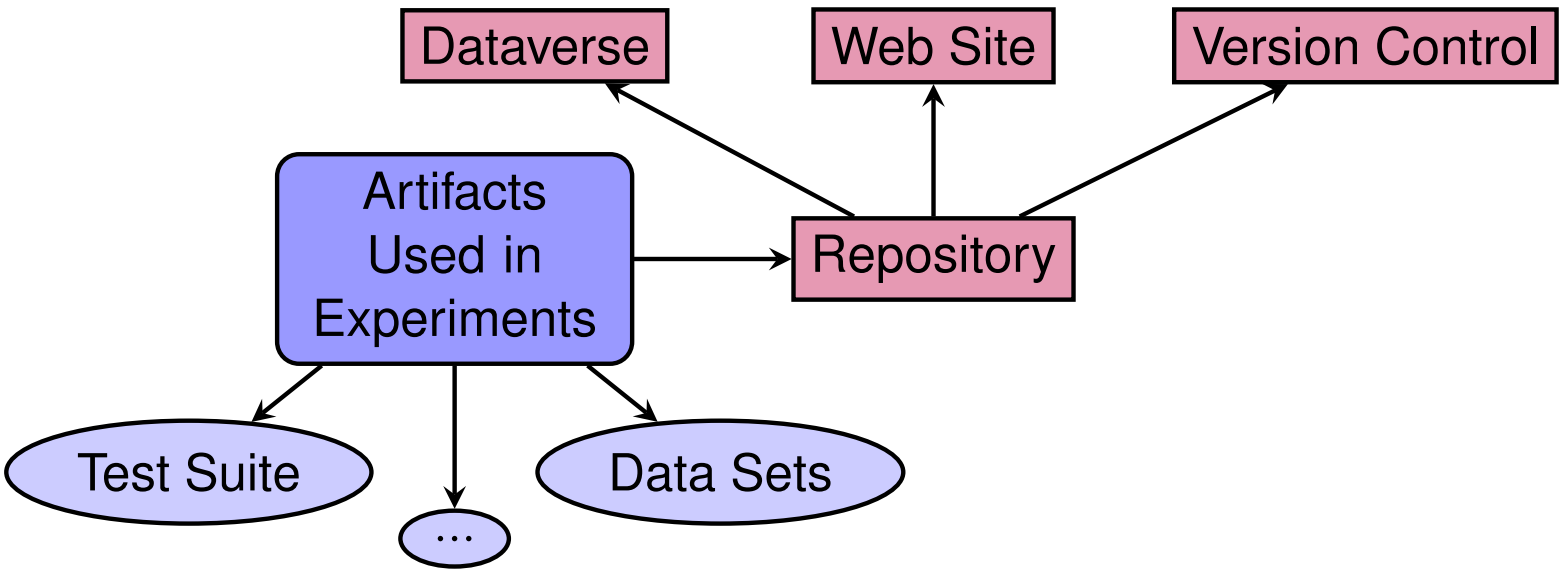


# Reproducible Research Through Sharing

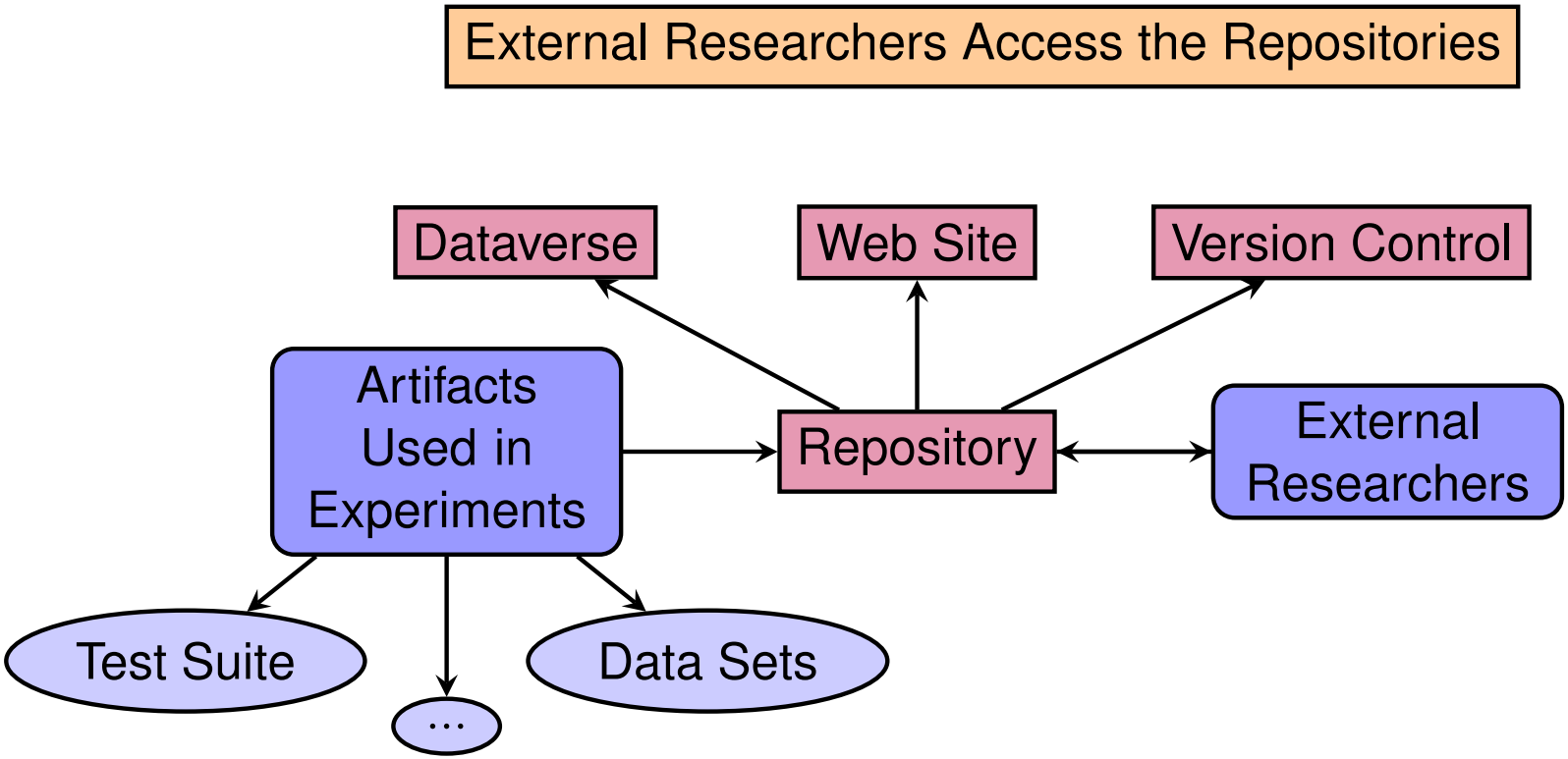


# Reproducible Research Through Sharing

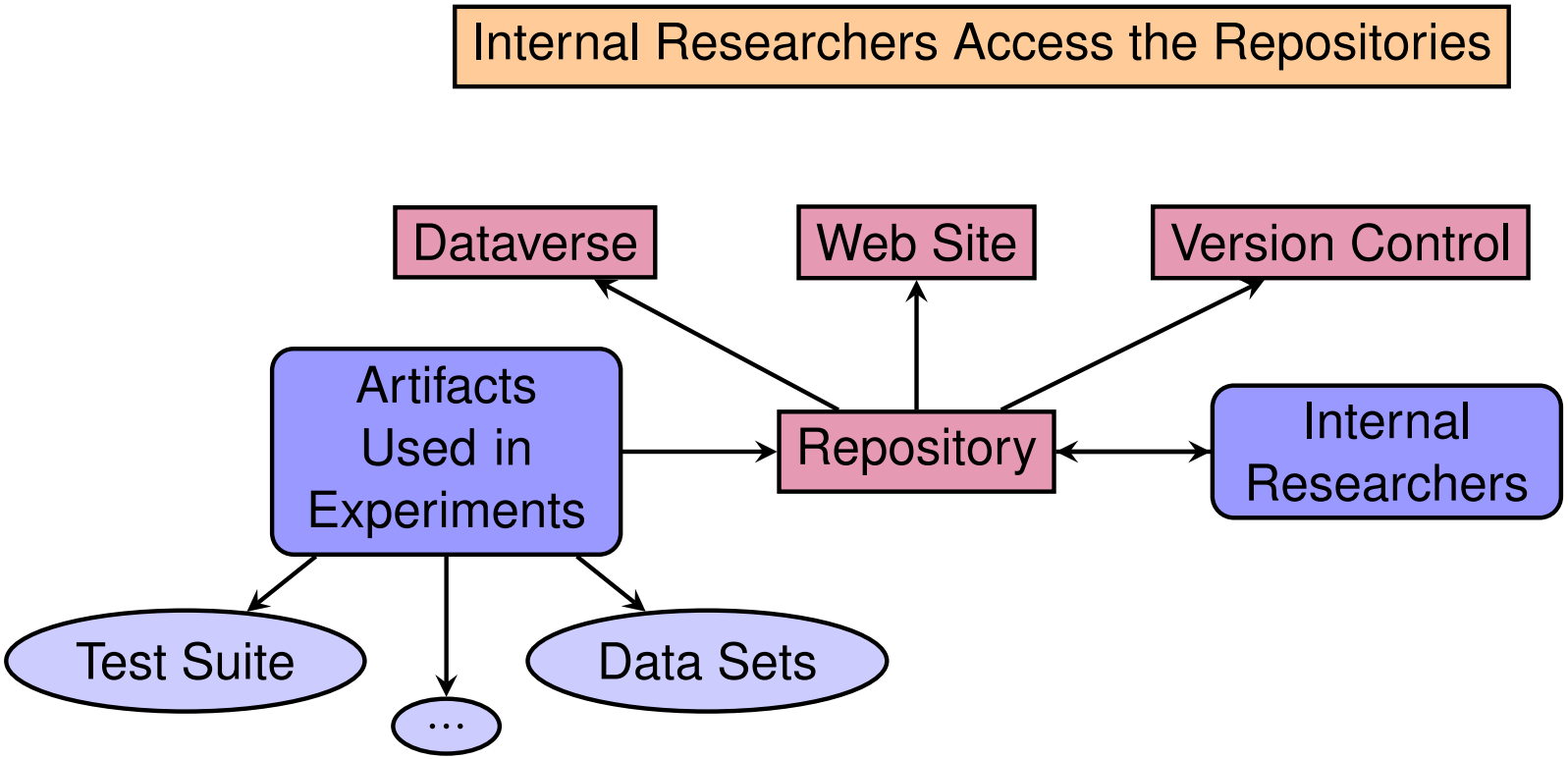
Deposit the Artifacts from Experimentation in One or More Repositories



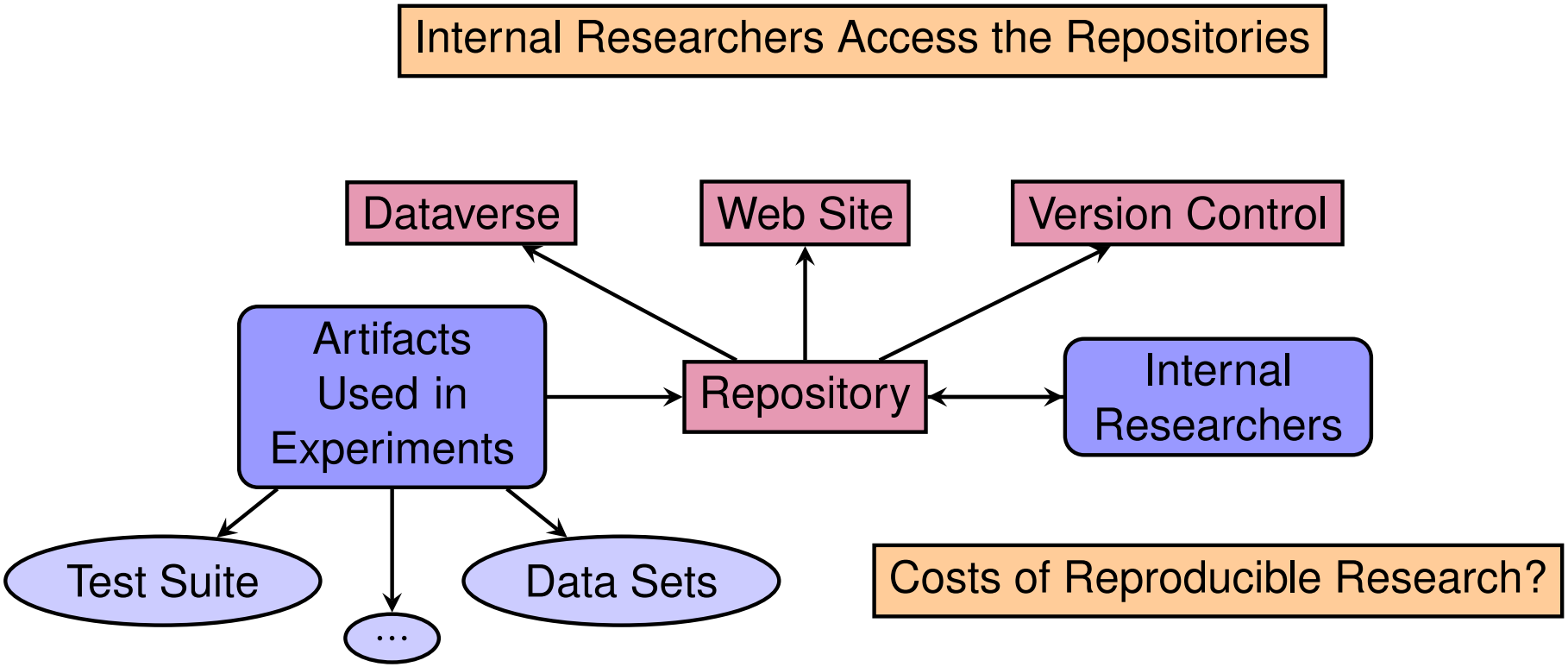
# Reproducible Research Through Sharing



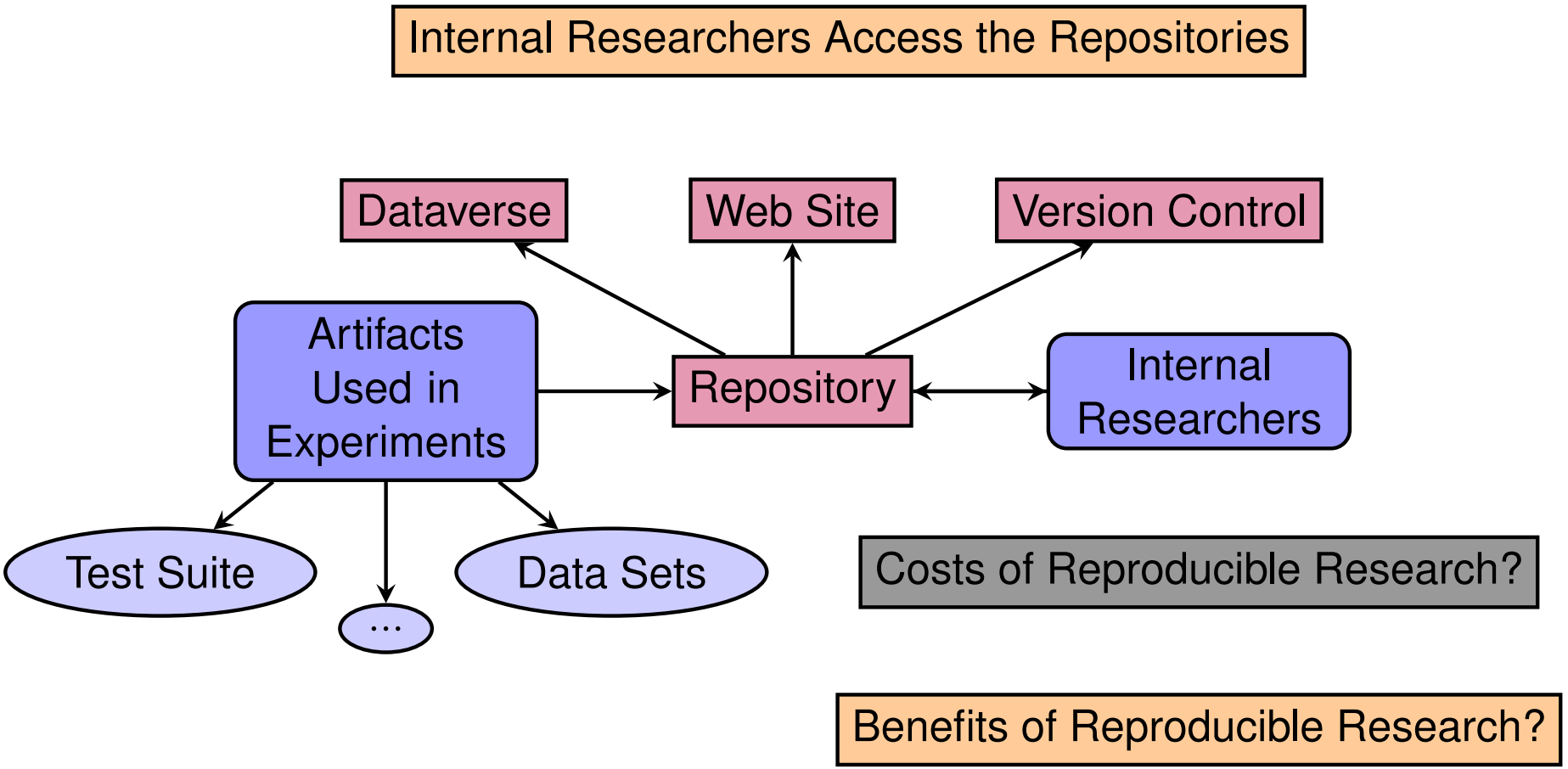
# Reproducible Research Through Sharing



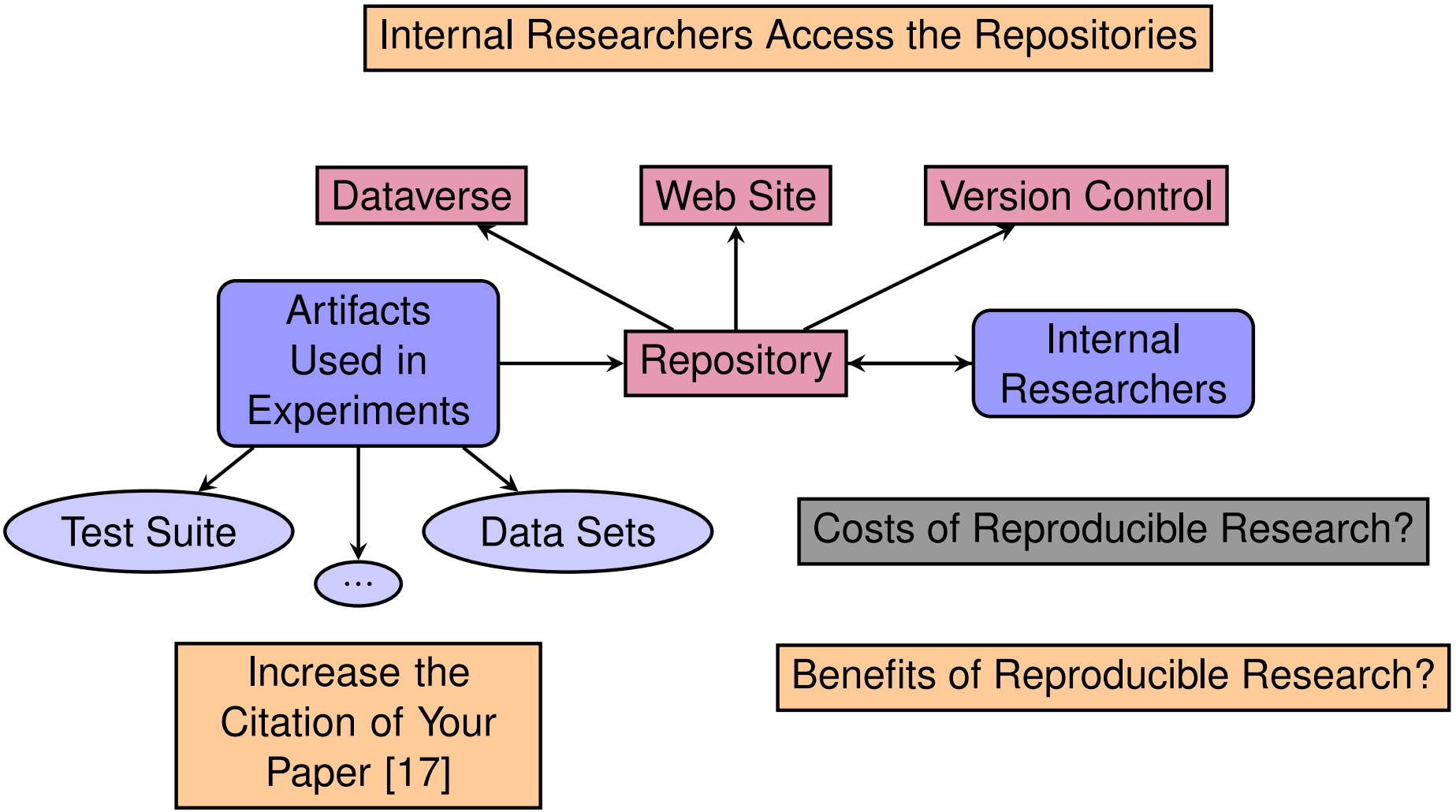
# Reproducible Research Through Sharing



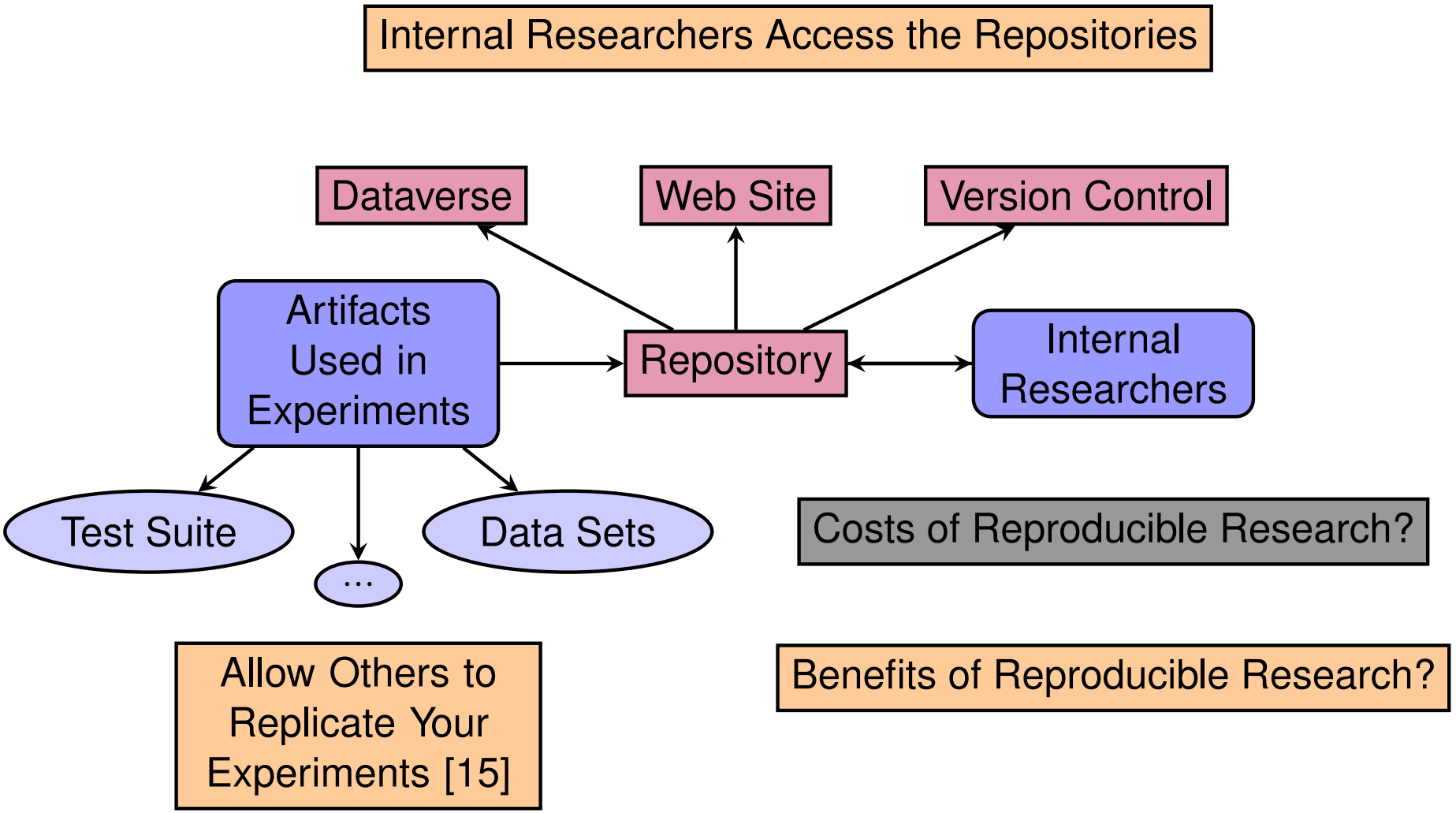
# Reproducible Research Through Sharing



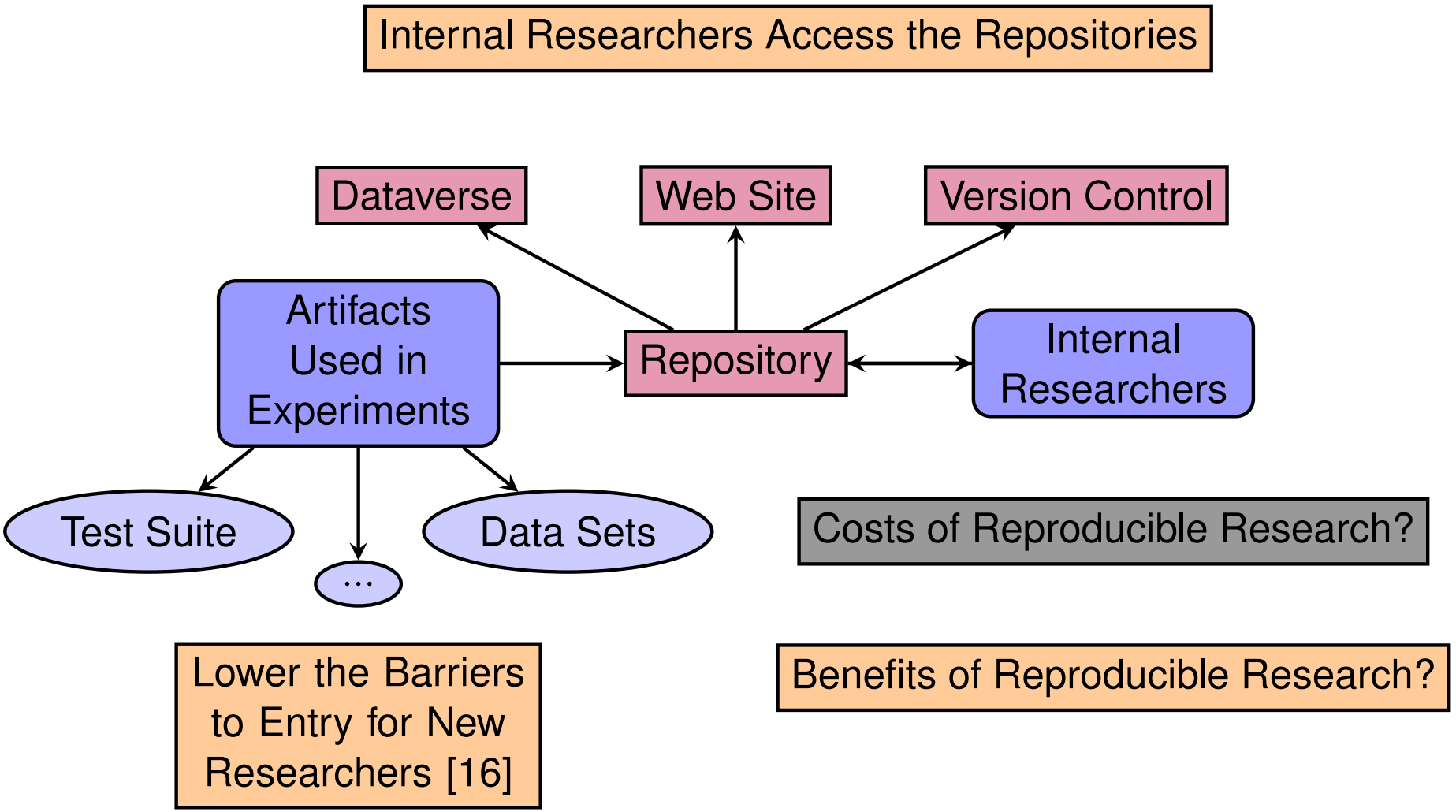
# Reproducible Research Through Sharing



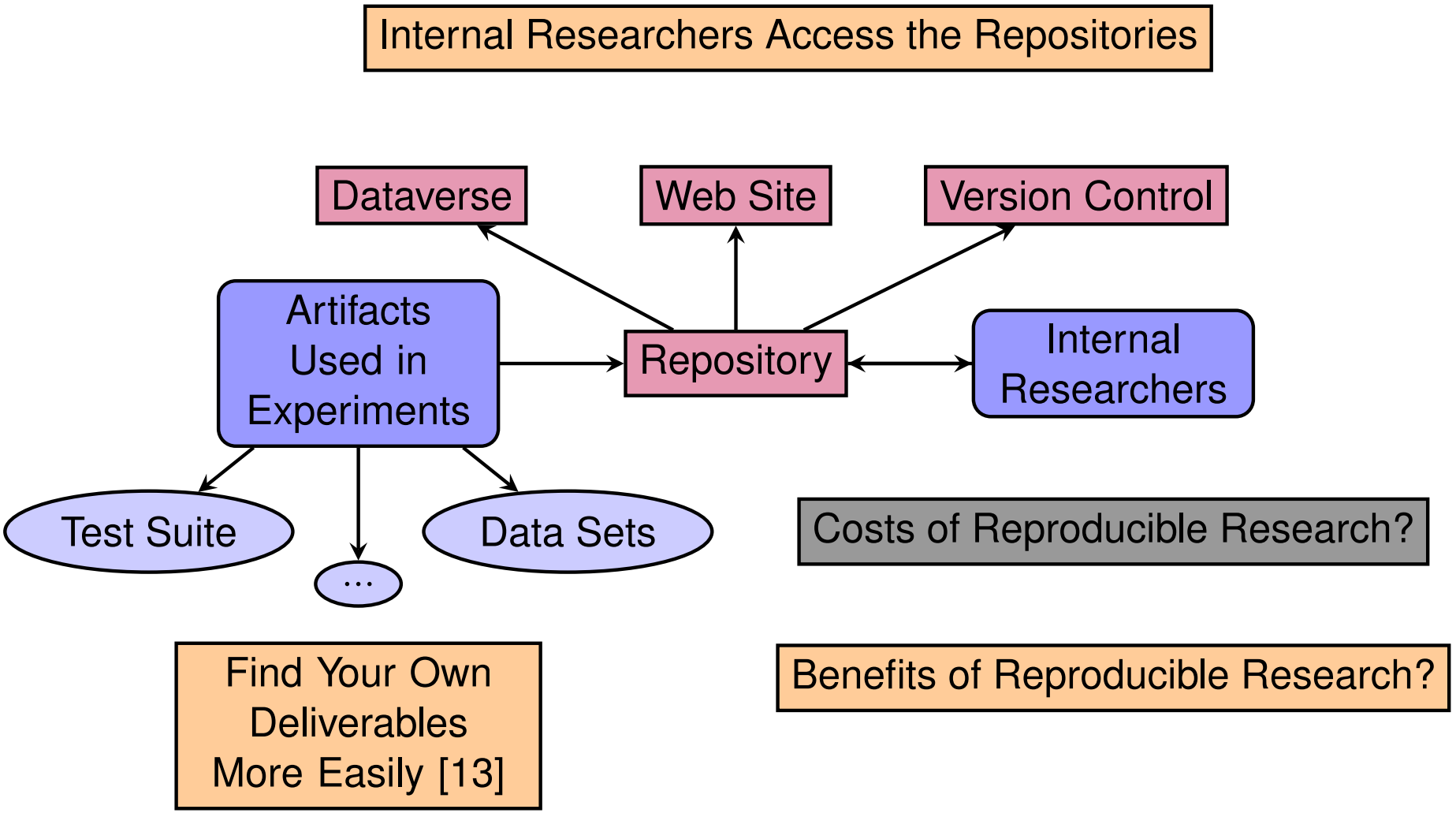
# Reproducible Research Through Sharing



# Reproducible Research Through Sharing



# Reproducible Research Through Sharing



# Practical Suggestions: Pick Tools Carefully

# Practical Suggestions: Pick Tools Carefully

Use the R Language for Statistical Computation When:

# Practical Suggestions: Pick Tools Carefully

Use the R Language for Statistical Computation When:

Conducting Experiments

# Practical Suggestions: Pick Tools Carefully

Use the R Language for Statistical Computation When:

- Conducting Experiments
- Visualizing Data

# Practical Suggestions: Pick Tools Carefully

Use the R Language for Statistical Computation When:

- Conducting Experiments
- Visualizing Data
- Data Mining

# Practical Suggestions: Pick Tools Carefully

Simple Shell or Full-Featured IDEs

Use the R Language for Statistical Computation When:

- Conducting Experiments
- Visualizing Data
- Data Mining

# Practical Suggestions: Pick Tools Carefully

Simple Shell or Full-Featured IDEs

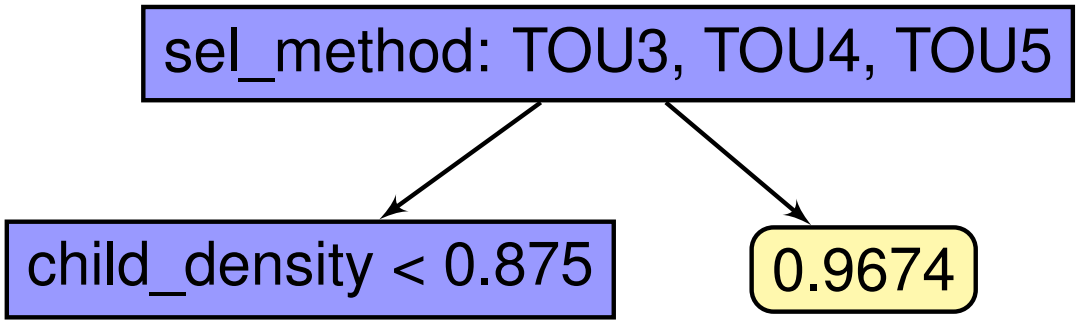
Use the R Language for Statistical Computation When:

- Conducting Experiments
- Visualizing Data
- Data Mining

Advanced Data Analysis Methods

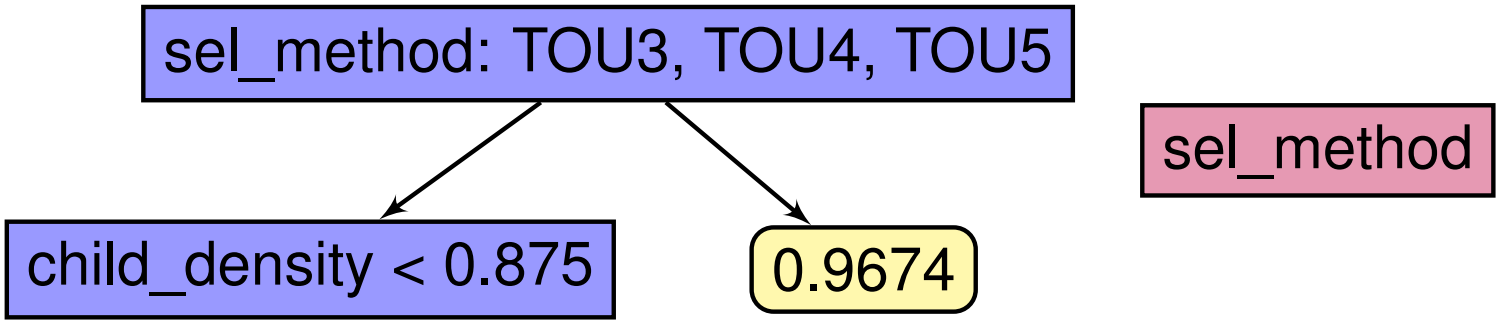
# Practical Suggestions: Use Tree Models

**RM**



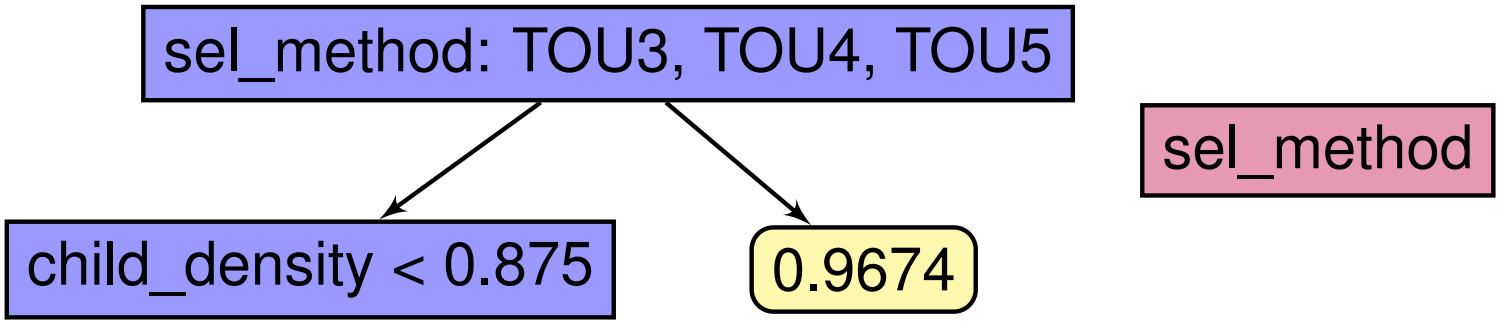
# Practical Suggestions: Use Tree Models

**RM**

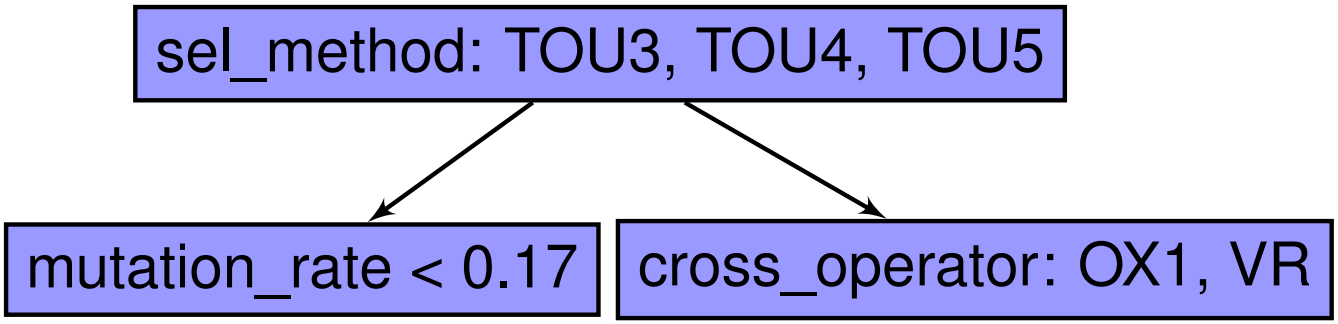


# Practical Suggestions: Use Tree Models

**RM**

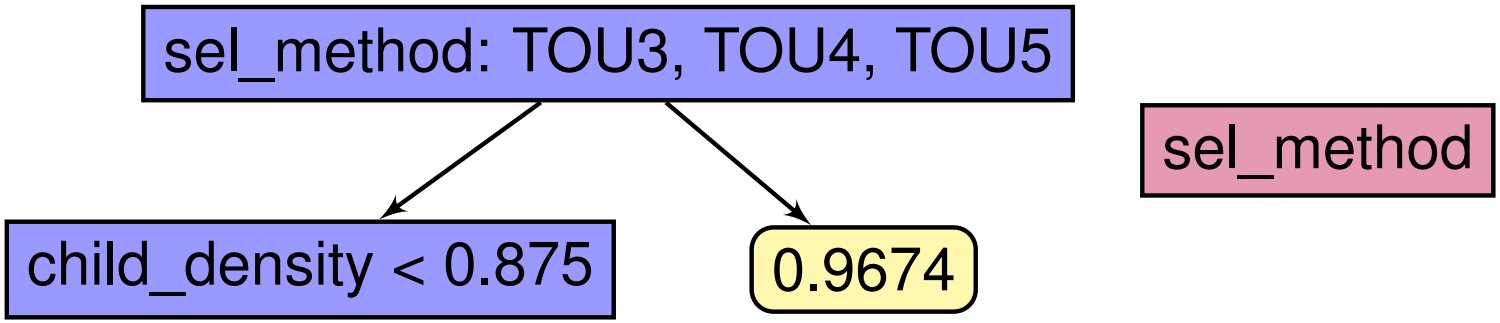



**GB**

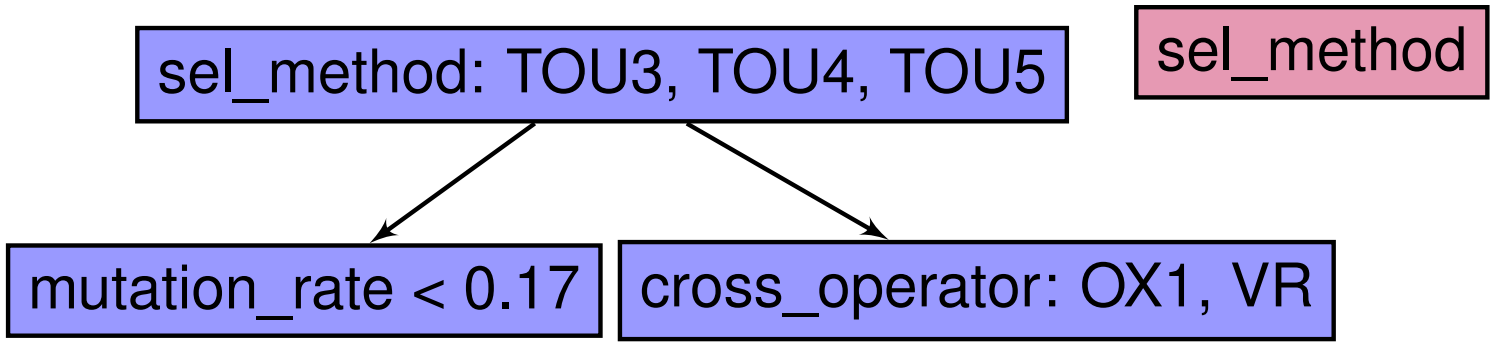


# Practical Suggestions: Use Tree Models

**RM**

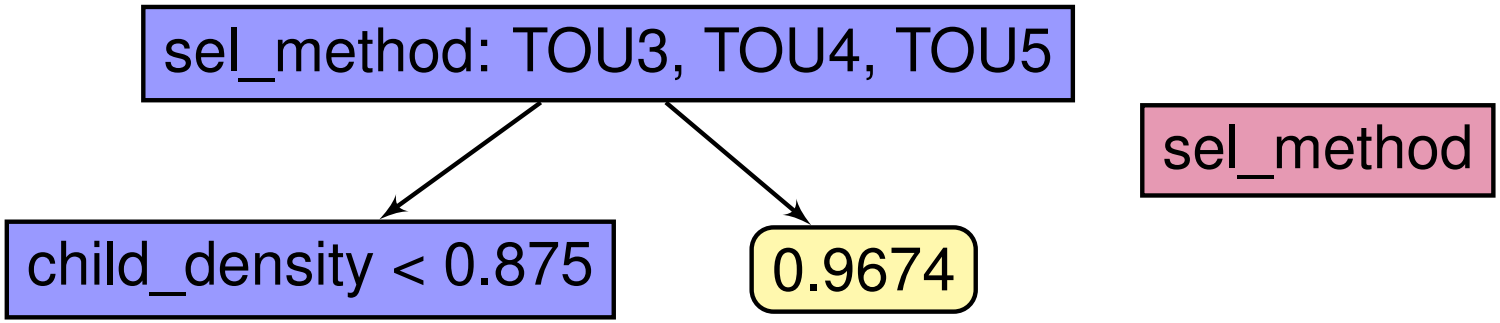



**GB**

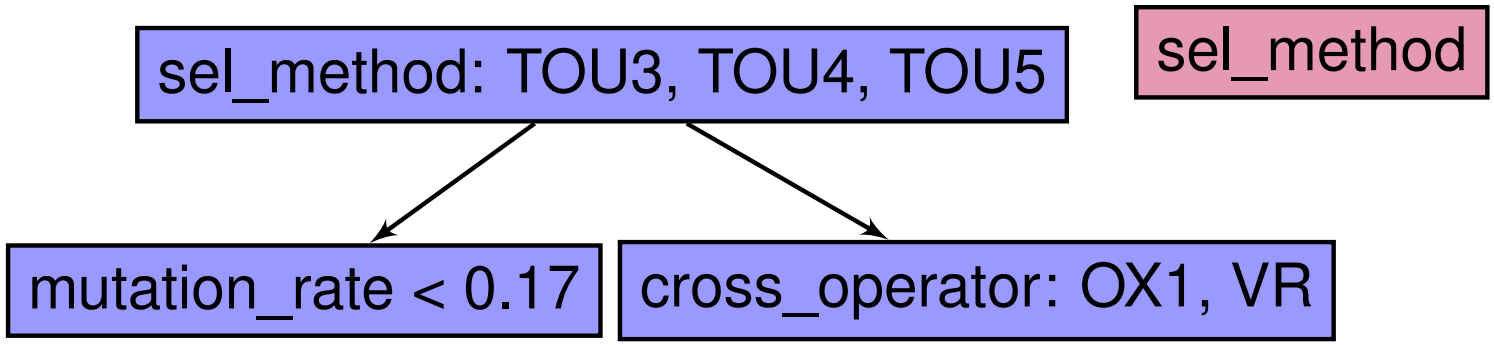


# Practical Suggestions: Use Tree Models

**RM**



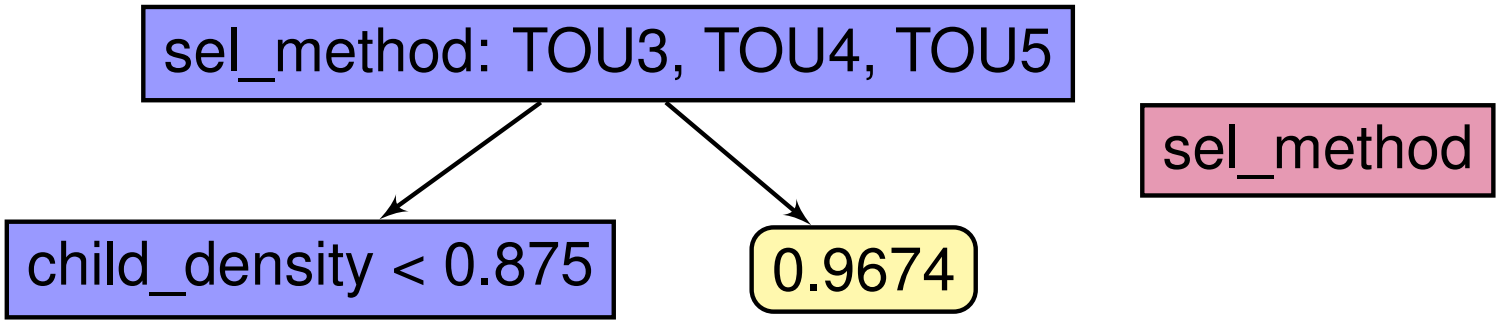

**GB**



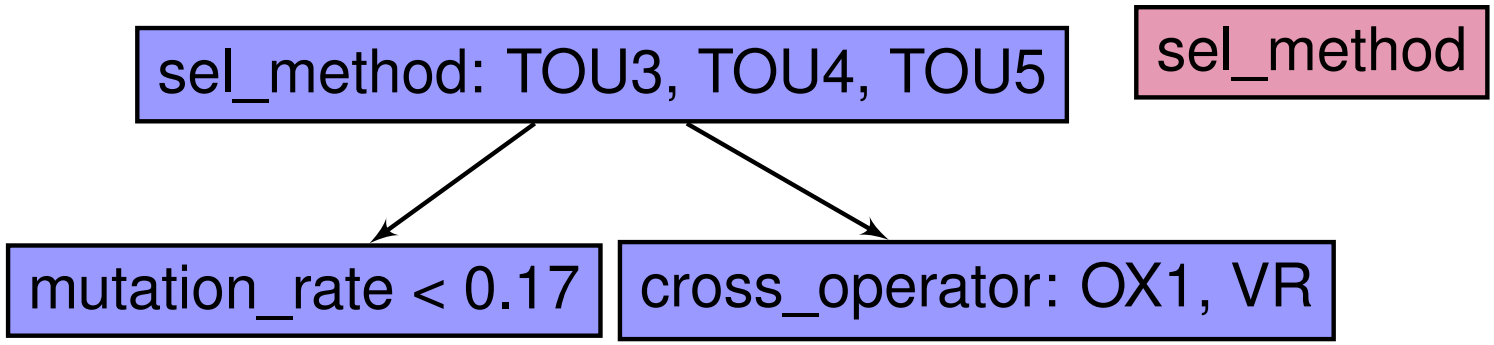

The sel_method variable is always the most important parameter

# Practical Suggestions: Use Tree Models

**RM**



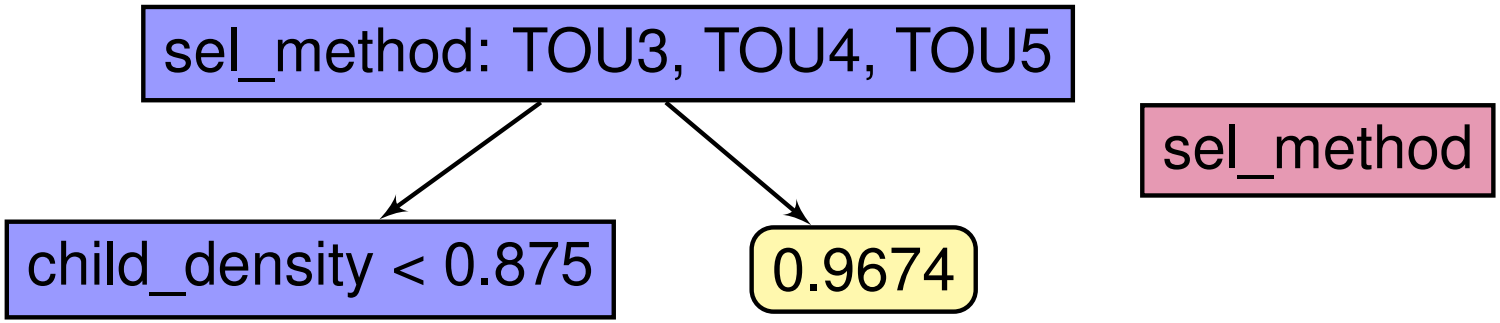

**GB**



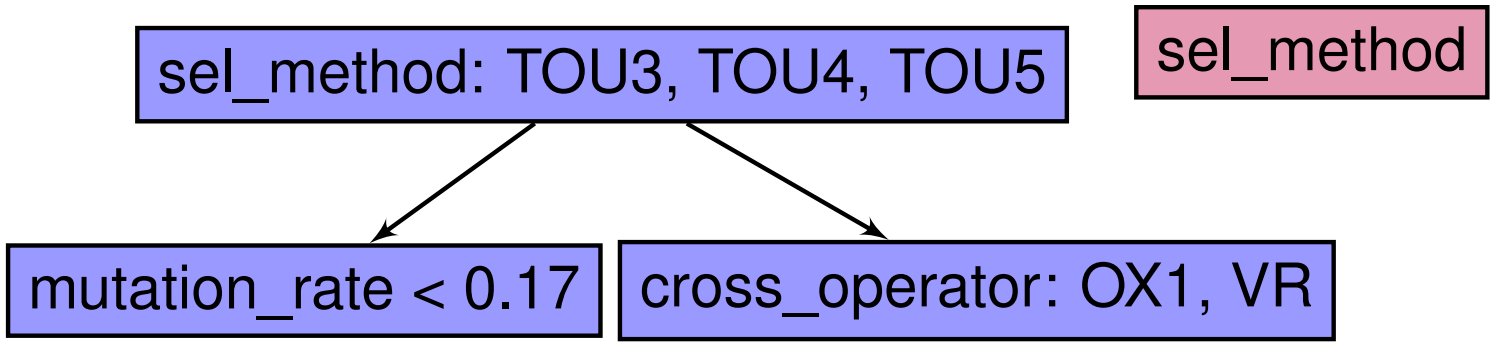

Importance of sel_method holds for all case study applications

# Practical Suggestions: Use Tree Models

**RM**



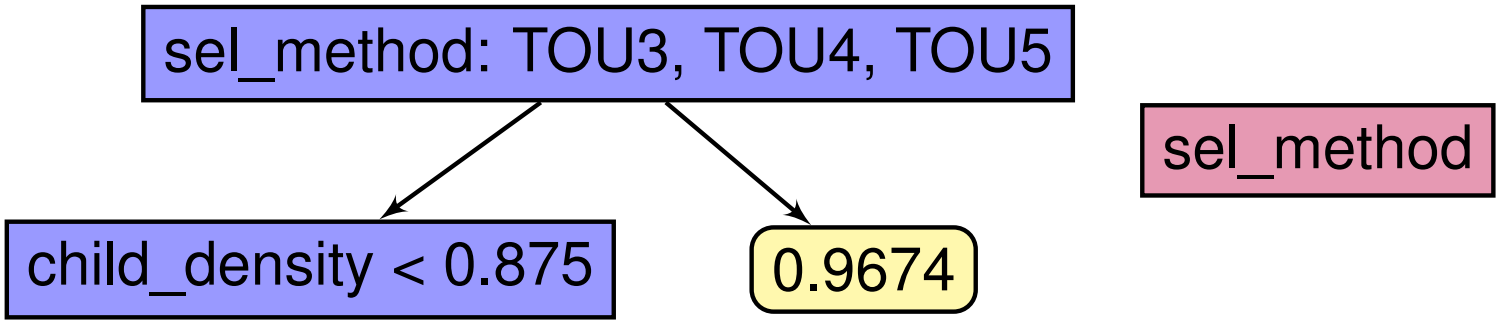

**GB**



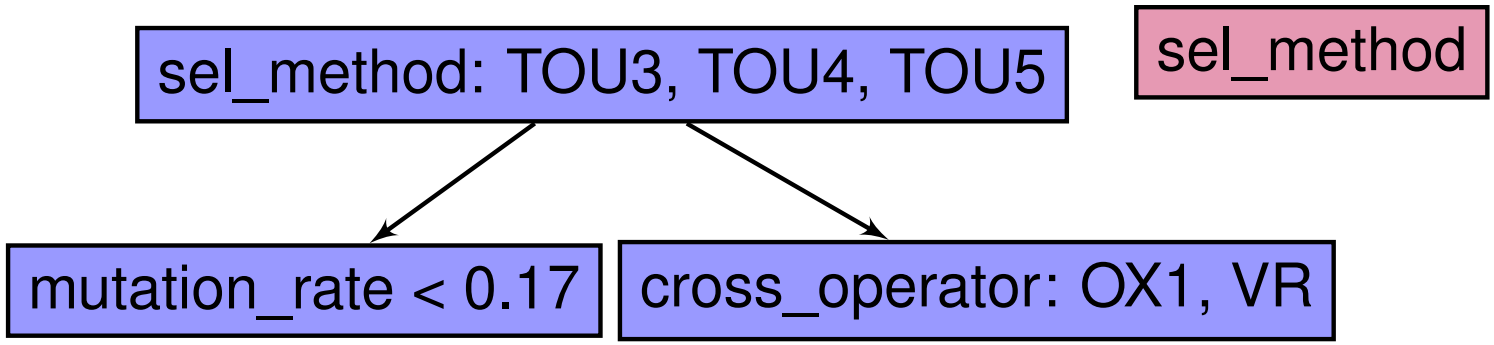

How does the selection method impact efficiency and effectiveness?

# Practical Suggestions: Use Tree Models

**RM**

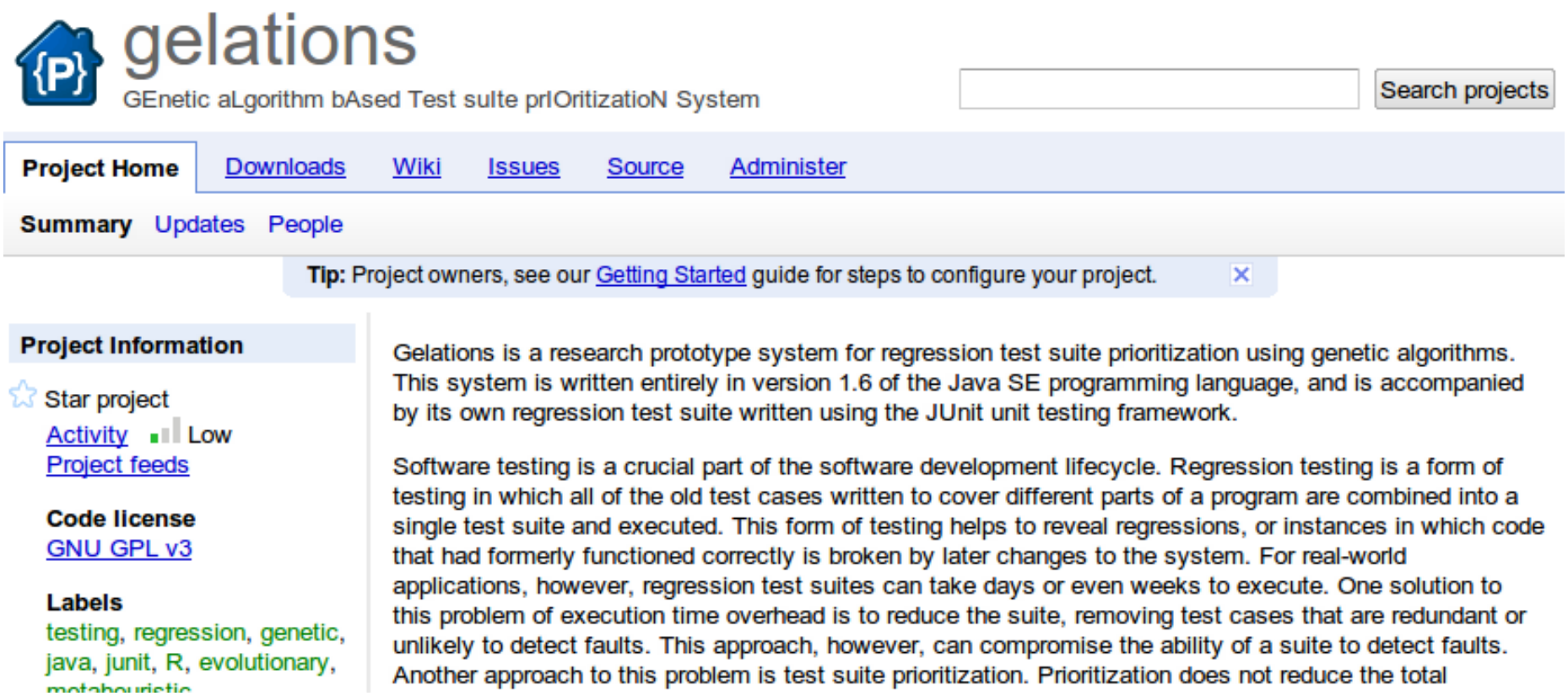



**GB**

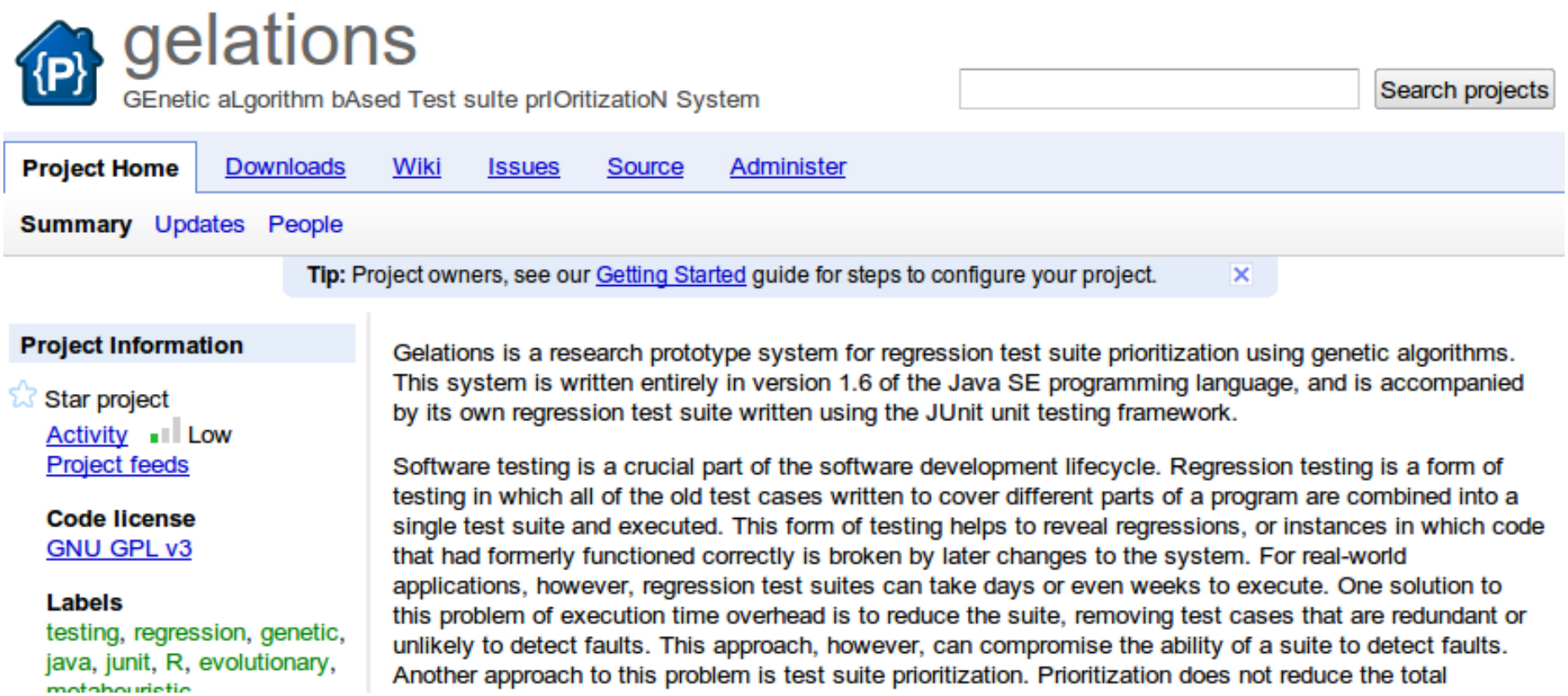



For More Details, Please See Conrad et al. [22]

# Practical Suggestions: Use Open Source

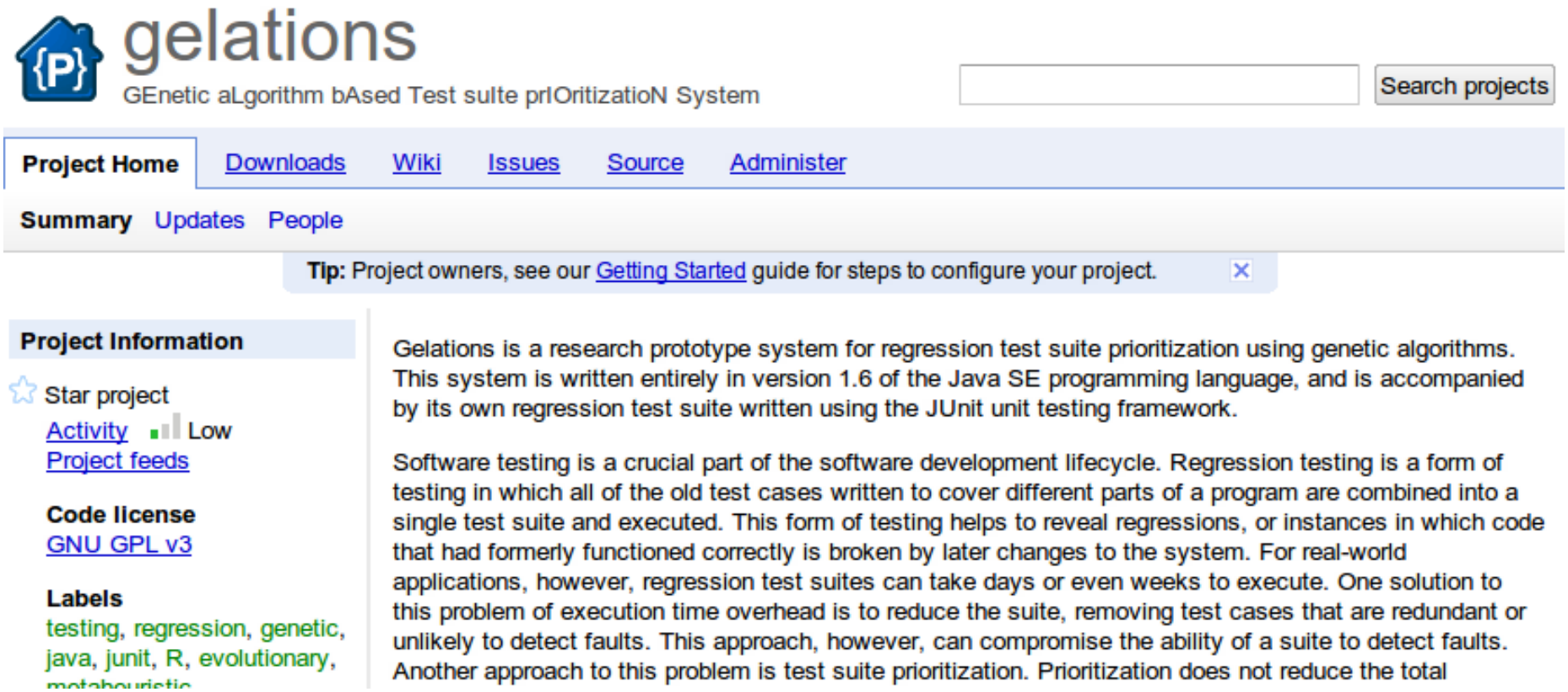


# Practical Suggestions: Use Open Source



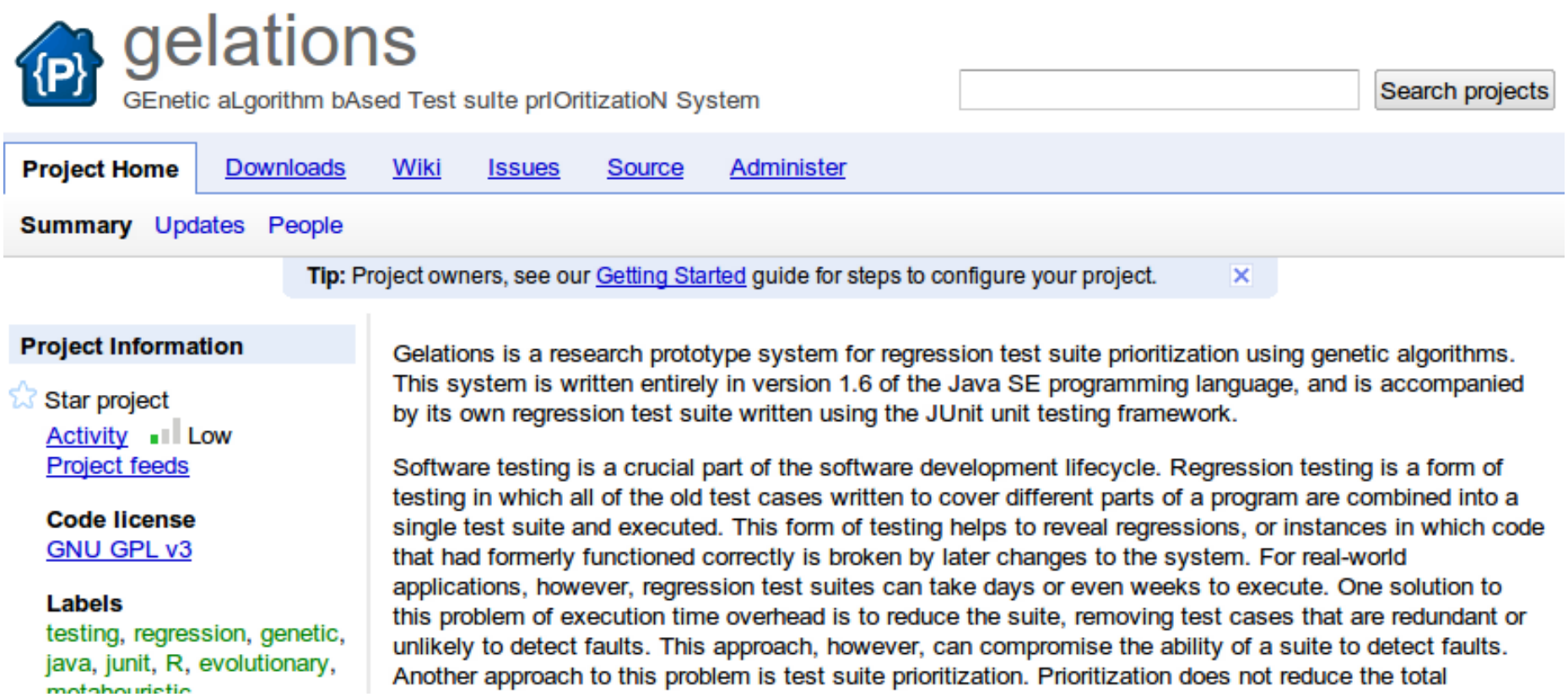

Gelations is a Genetic Algorithm-Based Test Suite Prioritizer

# Practical Suggestions: Use Open Source




Visit `http://gelations.googlecode.com/` for More Details

# Practical Suggestions: Use Open Source




Visit `http://raise.googlecode.com/` for Greedy Algorithms

# Conclusions and Future Work

The Potential Way Forward May Seem *Daunting*

# Conclusions and Future Work

The Potential Way Forward May Seem *Daunting*

Once you get your courage up and believe that you can do important problems, then you can.

Richard Hamming

# Conclusions and Future Work

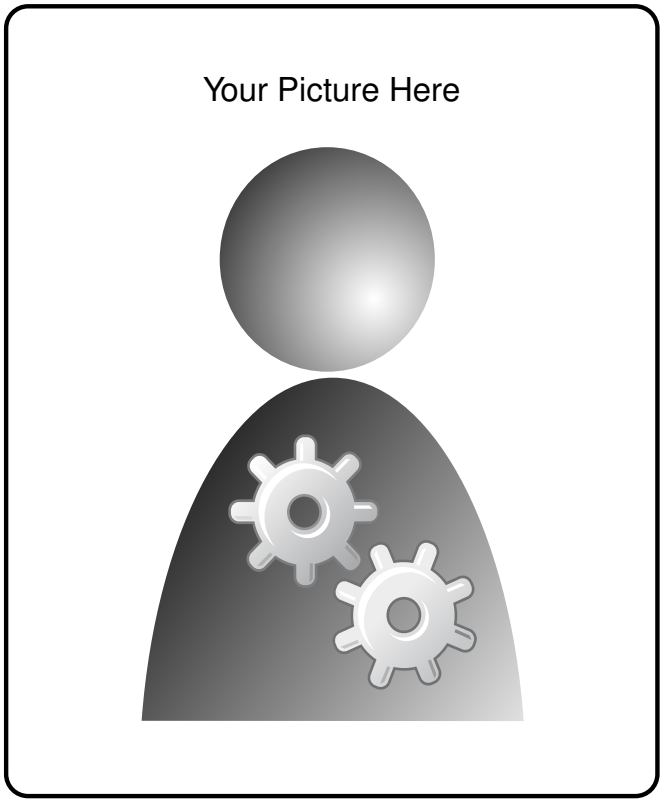
The Potential Way Forward May Seem *Daunting*

Once you get your courage up and believe that you can do important problems, then you can.

Richard Hamming

Share *One* Item With Your Next Paper

# Conclusions and Future Work



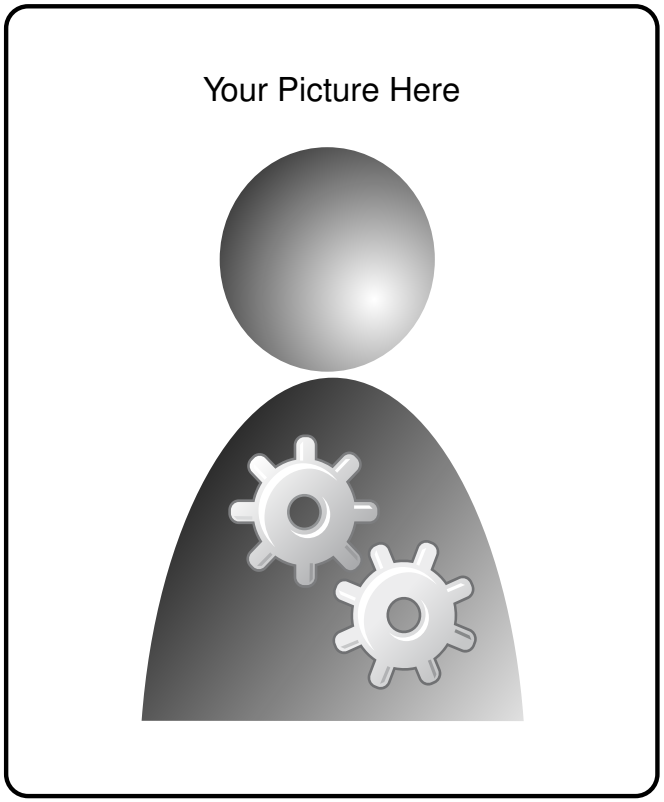

Leader of the Regression Testing Community

Your Help is Welcomed and Appreciated!

Once you get your courage up and believe that you can do important problems, then you can.

Richard Hamming

# Conclusions and Future Work



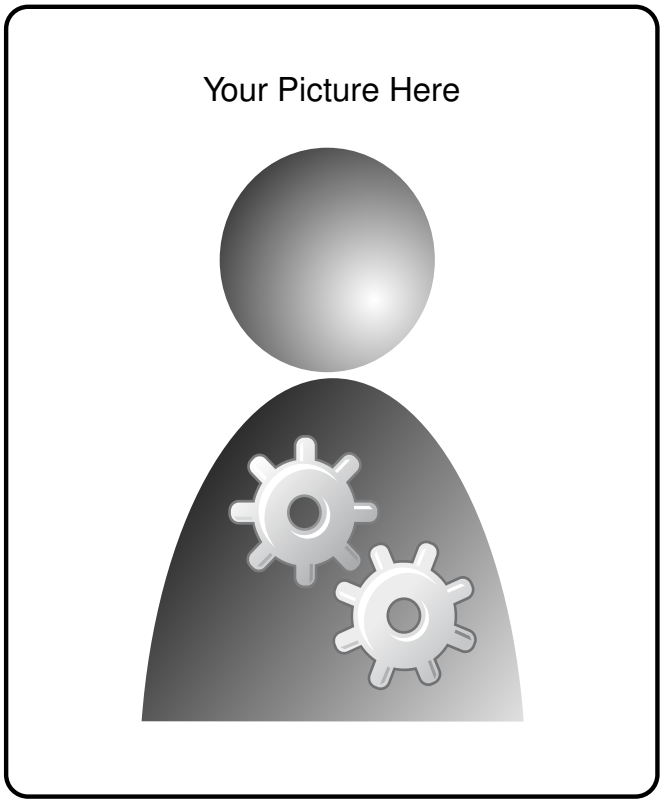

Leader of the Regression Testing Community

Read the Paper and Contact Me with Comments

Once you get your courage up and believe that you can do important problems, then you can.

Richard Hamming

# Conclusions and Future Work



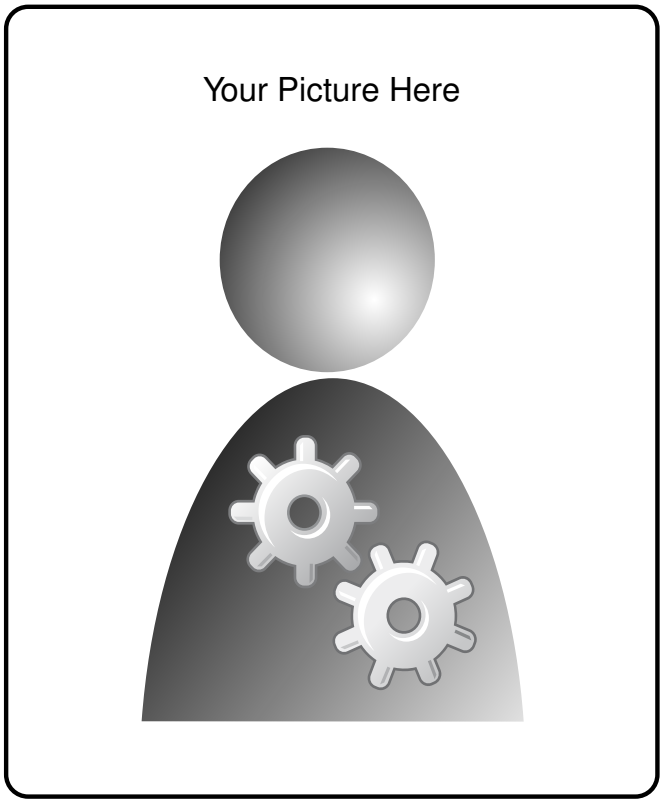

Leader of the Regression Testing Community

Read the Paper and Contact Me with Comments

**Future Work**:
Complete Case Study of Reproducible Research in Regression Testing

- ✓ Data Sets
- ✓ Tools
- ✓ Visualizations

# Conclusions and Future Work




Leader of the Regression Testing Community

Read the Paper and Contact Me with Comments

**Acknowledgments**:

- ✓ Neil Elliott
- ✓ Jonathan Kauffman
- ✓ Molly Mattis
- ✓ Zachary Williams

provided feedback on drafts of the paper